

Uniwersytet Warszawski
Wydział Biologii



Anna Miścicka
Nr albumu: 308952

Wykorzystanie niekanonicznych kodonów startu translacji do syntezy alternatywnych form białek kierowanych do mitochondriów

Rozprawa doktorska
w zakresie nauk biologicznych
dyscyplinie biologii

Praca wykonana pod kierunkiem
Prof. dr hab. Joanny Kufel
Instytut Genetyki i Biotechnologii Wydziału Biologii Uniwersytetu Warszawskiego

Warszawa, maj 2019

Oświadczenie kierującego pracą

Oświadczam, że niniejsza praca została przygotowana pod moim kierunkiem i stwierdzam, że spełnia ona warunki do przedstawienia jej w postępowaniu o nadanie stopnia doktora nauk biologicznych w zakresie biologii.

Data

Podpis kierującego pracą

Oświadczenie autora pracy

Świadoma odpowiedzialności prawnej oświadczam, że niniejsza rozprawa doktorska została napisana przeze mnie samodzielnie i nie zawiera treści uzyskanych w sposób niezgodny z obowiązującymi przepisami.

Oświadczam również, że przedstawiona praca nie była wcześniej przedmiotem procedur związanych z uzyskaniem stopnia doktora w innej jednostce.

Oświadczam ponadto, że niniejsza wersja pracy jest identyczna z załączoną wersją elektroniczną.

Data

Podpis autora pracy

Słowa kluczowe

Inicjacja translacji, non-AUG, mitochondrium, *Saccharomyces cerevisiae*, podwójna lokalizacja, alternatywne formy białek

Tytuł pracy w języku angielskim

Utilization of non-canonical translation start codons for the synthesis of alternative proteins variants targeted to mitochondria.

Streszczenie w języku polskim

Translacja jest jednym z najważniejszych procesów w każdej żywej komórce. U organizmów eukariotycznych zdecydowana większość transkryptów podlega kanonicznej translacji, która zachodzi według modelu skanującego. Po rozpoznaniu kodonu startu translacji przez małą podjednostkę rybosomu i przyłączeniu dużej podjednostki następuje synteza białka na matrycy mRNA, która ulega terminacji gdy rybosom dociera do kodonu STOP. Następuje wówczas dysocjacja rybosomu oraz odłączenie mRNA i nowopowstałego peptydu.

Inicjacja translacji nie zawsze jednak zachodzi z kodonu AUG. W niektórych przypadkach miejsce startu translacji stanowią również kodony inne niż AUG (non-AUG). Możliwe jest również wykorzystanie do translacji jednego transkryptu dwóch miejsc: kanonicznego kodonu AUG i alternatywnego non-AUG położonego poniżej lub powyżej. W wyniku takiego procesu syntetyzowanych jest kilka izoform białek, podstawowa powstała z kanonicznego miejsca inicjacji translacji i dodatkowe syntetyzowane z kodonu non-AUG. Wykorzystanie alternatywnego miejsca inicjacji translacji (aTIS) wydaje się bardziej powszechne u wirusów i w komórkach ssaczych, natomiast u drożdży *Saccharomyces cerevisiae* opisano do tej pory tylko pięć takich przypadków, dla których kanoniczny wariant białka jest obecny w cytozolu, a wydłużona izoforma lokalizuje się w mitochondrium.

W pracy wykazano, że u drożdży zjawisko syntezy dwóch form białkowych dzięki wykorzystaniu aTIS jest zdecydowanie bardziej powszechne niż sądzono i w wielu przypadkach prowadzi do podwójnej lokalizacji białek. Wykonane analizy bioinformatyczne wskazują, że około 2/3 białek w genomie drożdżowym może posiadać non-AUG aTIS, z czego ponad 500 zyskuje lokalizację mitochondrialną w formie wydłużonej. Eksperymentalnie potwierdzono lokalizację mitochondrialną 23 różnych białek, z których 19 nie było wcześniej obserwowane w mitochondriach. Ponadto, przeprowadzono systematyczne badania dotyczące podwójnej lokalizacji białka Trz1, które wykazały, że to właśnie wydłużona izoforma odpowiada za jego funkcje mitochondrialne.

Dane przedstawione w pracy pozwalają na wyjaśnienie złożoności i różnorodności proteomu mitochondrialnego oraz wskazują na znaczący wpływ inicjacji translacji z kodonów non-AUG w tworzeniu alternatywnych form białek, które mogą pełnić różnorakie funkcje.

Streszczenie w języku angielskim

Translation is one of the most important processes in every living cell. In eukaryotes the majority of transcripts are translated in a canonical manner using the scanning model of translation. After joining of the large ribosome subunit, the mature ribosome starts translation and synthesizes protein until the STOP codon is reached. The ribosome then splits into two subunits and the mRNA which encoded the nascent protein is released.

However, translation initiation does not always occur at an AUG start codon. In some cases, non-AUG codons serve as translation initiation start sites. It is possible to utilise two different codons as start sites in one transcript: a canonical AUG and an alternative upstream or downstream non-AUG. Utilization of non-AUG alternative translation initiation sites (aTIS) is most common in viruses and mammalian cells. In the yeast *Saccharomyces cerevisiae*, five described cases concern proteins that are translated from upstream near-cognate start codons as N-terminally extended variants that localize to mitochondria.

In this work I show that in yeast the potential for producing alternative protein isoforms by non-AUG translation initiation is much more prevalent than previously anticipated and may generate an unidentified pool of mitochondrial proteins. A bioinformatics analysis performed in my research group shows that about 2/3 of proteins in the yeast genome have predicted one or more upstream aTIS, with more than 500 gains in mitochondrial localisation as the extended isoforms. I confirmed mitochondrial localisation of 23 different proteins previously not recognized as mitochondrial whose standard forms do not carry an MTS. I also analysed the dual localisation of the Trz1 protein, which has been shown to localise and function in both the nucleus and mitochondria. I verified that the extended non-AUG isoform is responsible for mitochondrial localisation and functions of Trz1.

My data highlight the potential of non-canonical translation initiation in expanding the capacity of the mitochondrial proteome and possibly also other cellular features.

Dziękuję:

*mojej promotorce, prof. dr hab. Joannie Kufel za możliwość wykonywania pracy doktorskiej
w jej grupie, opiekę merytoryczną oraz umożliwienie mi obrony w terminie.*

*mgr Lounisowi Zenad i mgr Michałowi Świrskiemu za przeprowadzenie
badań bioinformatycznych, bez których nie mogłabym wykonać mojego projektu.*

*wszystkim członkom grupy badawczej prof. Joanny Kufel za wspólnie spędzony czas oraz wiedzę
i umiejętności, których nauczyłam się od nich podczas doktoratu,
a w szczególności dziękuję mgr Agnieszce Czarnockiej-Cieciurze
za nieocenioną pomoc, wsparcie i wspólne rozmowy każdego dnia.*

*moim studentkom, Natalii Wołkowskiej oraz Kasi Iwańskiej za uśmiech, zaangażowanie
w projekt (i wykonywanie rzeczy, których mi się nie chciało, takich jak wylewanie
pożywek) oraz to wszystko, czego wzajemnie się od siebie nauczyłyśmy.*

*I would like to thank prof. Gretchen Edwalds-Gilbert for her priceless help
and being one of the most positive person I ever met.*

*pracownikom, doktorantom i studentom Instytutu Genetyki i Biotechnologii
za to wszystko, co czyni nasz Instytut unikatowym.*

*za nieocenioną pomoc merytoryczną:
prof. dr hab. Pawłowi Golikowi, dr Karolinie Łabędzkiej-Dmoch (IGiB)
prof.dr hab. Agnieszce Chacińskiej i dr Lidii Wróbel (IIMCB)*

mojej rodzinie i przyjaciołom – wiedzę za co.

Spis treści

Oświadczenie kierującego pracą	2
Oświadczenie autora pracy.....	2
Słowa kluczowe.....	3
Tytuł pracy w języku angielskim	3
Streszczenie języku polskim	4
Streszczenie w języku angielskim.....	5
Wykaz skrótów, najczęściej używanych w pracy:	9
Wstęp	10
1.Wprowadzenie	10
2.Kanoniczna translacja	11
3.Niekanoniczna translacja	14
3.1.Niekanoniczna inicjacja translacji.....	14
3.1.1.Mechanizmy translacji niezależne od struktury kapu 5'	14
3.1.2.Mechanizmy inicjacji zależne i niezależne od skanowania mRNA.....	17
3.1.3.Inicjacja translacji z kodonów niekanonicznych non-AUG.....	21
3.2.Niekanoniczna elongacja i terminacja translacji (rekodowanie).....	24
4.Regulacja wyboru niekanonicznego kodonu startu	29
5.Mitochondrium i jego funkcje u drożdży <i>Saccharomyces cerevisiae</i>	32
Cel pracy	35
Materiały i metody	36
1.Część bioinformatyczna.....	36
1.1.Wykrywanie wydłużonych form białek	36
1.2.Analiza GO Term (Gene Ontology).....	36
1.3.Określenie lokalizacji komórkowej	36
1.4.Analiza konserwacji	37
2.Część molekularna	38
2.1.Podstawowe techniki biologii molekularnej	38
2.2.Hodowla drożdży	38
2.3.Konstrukcja plazmidów.....	39
2.4.Transformacja drożdży.....	41
2.5.Sporulacja i rozkładanie tetrad	42
2.6.Fracjonowanie komórek drożdżowych	42

2.7.Precypitacja białek przy użyciu kwasu trójchlorooctowego (TCA).....	43
2.8.Deglikozylacja endoglikozydazą H (EndoH)	43
2.9.Defosforylacja białkową fosfatazą faga Lambda (LPP).....	44
2.10.Elektroforeza białek SDS-PAGE	44
2.11.Analiza western blot	45
2.12.Ekstrakcja małych RNA z drożdży	45
2.13.Analiza northern blot.....	46
Wyniki	48
1.Wykrywanie potencjalnych niekanonicznych startów translacji (aTIS).....	48
2.Identyfikacja MTS w wydłużonych na końcu N izoformach	53
3.Lokalizacja białek NTE w mitochondriach	58
4.Potwierdzenie funkcji mitochondrialnych białek MT-NTE	64
5.Konserwacja NTE u wyższych eukariontów	72
6.Podsumowanie	75
Dyskusja	76
1.Synteza białek NTE zwiększa różnorodność proteomu.....	76
2.Czynniki wpływające na wybór aTIS u drożdży	77
3.W wyniku inicjacja translacji z kodonów startu non-AUG powstaje wiele białek o lokalizacji mitochondrialnej.....	78
4.Funkcje mitochondrialnych izoform białkowych	81
Literatura	84

Wykaz skrótów, najczęściej używanych w pracy:

aa-tRNA – aminoacylowany tRNA

ABCE1 - białko wiążące kasetę ATP z podrodziny E (ang. *ATP-binding cassette subfamily E member 1*;

aTIS – alternatywne miejsce inicjacji translacji (ang. *alternative translation initiation site*)

CITE – wzmacniacze translacji niezależne od struktury kapu (ang. *cap-independent translation enhancers*)

eEF – eukariotyczny czynnik elongujący (ang. *eukariotic elongation factor*)

eIF – eukariotyczny czynnik inicjujący (ang. *eukariotic initiation factor*)

eRF – eukariotyczny czynnik uwalniający (ang. *eukaryotic release factor*)

IRES – wewnętrzne miejsce wiązania rybosomu (ang. *internal ribosome entry site*)

Miejsca E, P, A – miejsca w rybosomie, kluczowe dla mechanizmu translacji: wyjścia E (ang. *exit site*),
peptydylowe P (ang. *peptidyl site*) i aminoacylowe A (ang. *aminoacyl site*)

MT-NTE – mitochondrialne białko NTE

MTS⁻ – izoformy NTE tracące lokalizację mitochondrialną w stosunku do formy podstawowej (ang. *MTS loss*)

MTS – sygnał lokalizacji mitochondrialnej (ang. *mitochondrial targeting signal*)

MTS⁺ – izoformy NTE zyskujące lokalizację mitochondrialną w stosunku do formy podstawowej (ang. *MTS gain*)

Non-AUG – niekanoniczny kodon startu translacji (inny niż AUG)

NTE – wydłużenie na końcu N (ang. *N-terminal extension*)

ORF – otwarta ramka odczytu (ang. *open reading frame*)

PABP – białko wiążące ogon poli(A) (ang. *poly(A) binding protein*)

PIC – kompleks preinicjujący (ang. *preinitiation complex*)

PRF – programowana zmiana ramki odczytu (ang. *programmed ribosomal frameshifting*)

Ribo-seq – analiza profilowania rybosomów (ang. *ribosome profiling*)

SGD – Saccharomyces Genome Database

SLIC – metoda jednoetapowego klonowania niezależnego od ligacji (ang. *one-step sequence- and ligation-independent cloning*)

TAP(-tag) – znacznik TAP (ang. *Tandem Affinity Purification tag*)

TIS – miejsce inicjacji translacji (ang. *translation initiation site*) tu: miejsce inicjacji translacji z kanonicznego kodonu AUG

TISU – inicjator translacji krótkich 5'UTR (ang. *translation initiator of short 5' UTRs*)

TSS – miejsce startu transkrypcji (ang. *transcription start site*)

TURBS - terminacja powyżej miejsca wiązania rybosomu (ang. *termination upstream ribosome binding site*)

uORF – otwarta ramka odczytu, położona powyżej (ang. *upstream ORF*)

UTR 5' – rejon 5' niepodlegający translacji (ang. *5' untranslated region*)

YPD – pożywka pełna do hodowli drożdży z glukozą jako źródłem węgla

YPG – pożywka pełna do hodowli drożdży z glicerolem jako źródłem węgla

YPGal – pożywka pełna do hodowli drożdży z galaktozą jako źródłem węgla

Wstęp

1. Wprowadzenie

Translacja (z łaciny *translatio*) jest jednym z podstawowych procesów, zachodzących w każdej komórce. Jest ostatnim etapem transferu genów według centralnego dogmatu biologii molekularnej (Crick, 1970), a jednocześnie najbardziej kosztownym energetycznie procesem w komórce (Gunišová et al., 2018). Składa się z trzech etapów: inicjacji, elongacji oraz terminacji, które są regulowane przez odpowiednie czynniki translacyjne.

Kanoniczne miejsce startu translacji stanowi kodon AUG, który odpowiada za inkorporację metioniny u organizmów eukariotycznych lub formylometioniny u bakterii. Standardowy kod genetyczny obejmuje trzy kodony STOP: UAA (*orche*), UAG (*amber*), UGA (*opal*), lecz u niektórych glonów i pierwotniaków może być ich mniej (Jackson et al., 2012).

Synteza białka odbywa się przy udziale skomplikowanej maszyneryi translacyjnej, której głównym składnikiem są rybosomy: kompleksy rybonukleoproteinowe zbudowane z RNA i białek. Rybosomy zostały odkryte w 1955 roku, kiedy zaobserwowano je pod mikroskopem jako małe granule w cytoplazmie, składające się głównie z RNA (Palade, 1955). Rybosomalny RNA (rRNA) stanowi przeważającą większość całkowitego RNA w komórce. U eukariontów składają się z dwóch podjednostek: małej 40S, która odpowiada za wiązanie mRNA i dużej 60S, gdzie powstaje łańcuch polipeptydowy. Duża podjednostka posiada trzy miejsca, biorące udział w biosyntezie białka: miejsce wyjścia E (ang. *exit site*), peptydylowe P (ang. *peptidyl site*) i aminoacylowe A (ang. *aminoacyl site*) (Ramakrishnan, 2002). Ze względu na szczególnie istotną rolę rybosomów w komórce ich budowa, włączając strukturę poszczególnych miejsc i ich położenie w rybosomie, została dokładnie poznana dzięki badaniom krystalograficznym oraz obserwacjom mikroskopowym zamrożonych preparatów techniką Cryo-EM w rozdzielczości 2,9 Å dla najbardziej stabilnych rejonów (Ramakrishnan, 2014; Brown i Shao, 2018).

Kanoniczna translacja

Zdecydowana większość eukariotycznych mRNA jest syntetyzowana w sposób kanoniczny, zależny od struktury kapu 5'. Pierwszym procesem translacji jest inicjacja, polegająca na przyłączeniu się rybosomu do matrycowego RNA (mRNA) i inkorporacji pierwszego aminokwasu.

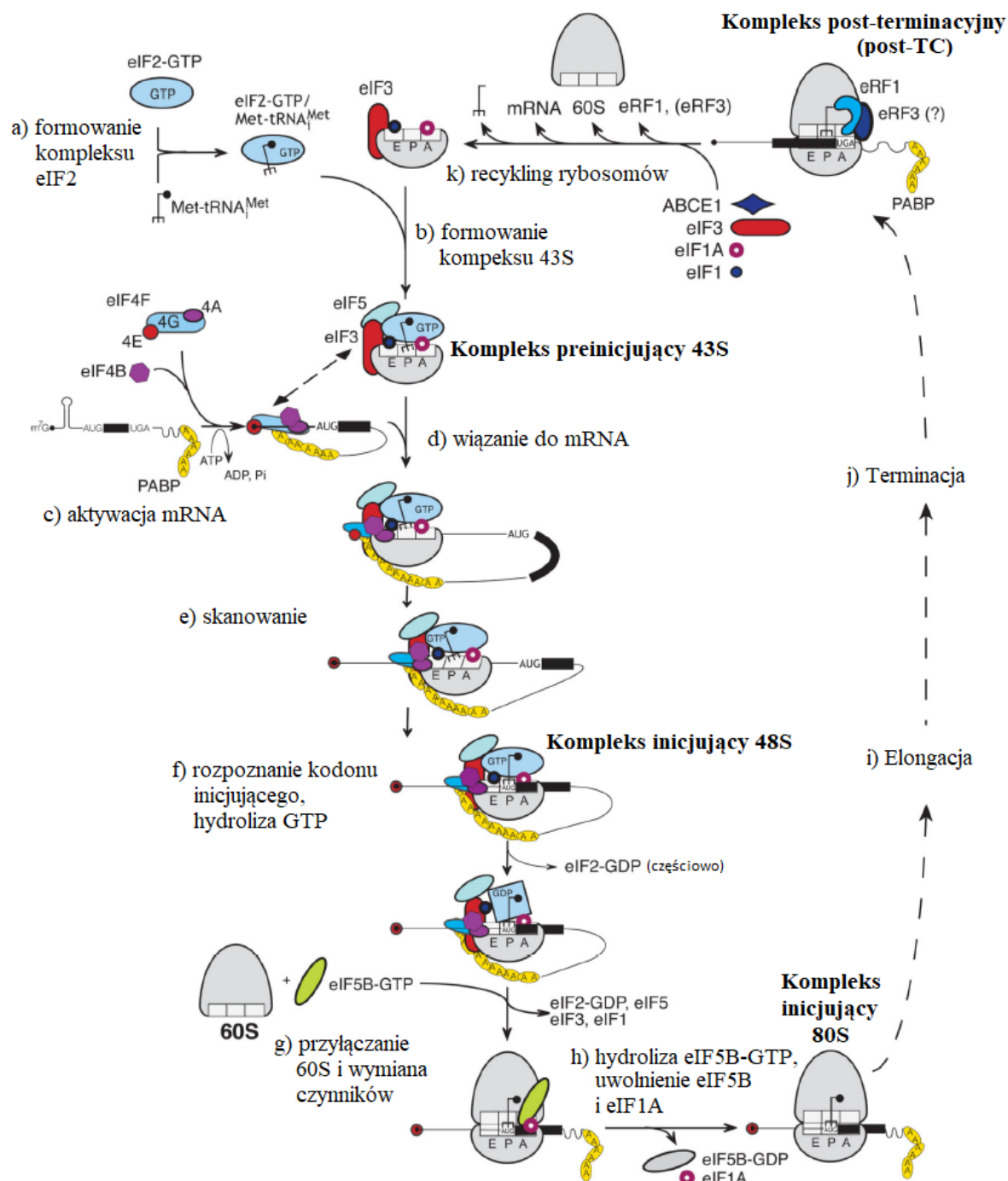


Fig 1. Model kanonicznej translacji zależnej od struktury kapu 5' u organizmów eukariotycznych. Na schemacie położono szczególny nacisk na inicjację translacji (a-h). Na podstawie (Jackson et al., 2010). Szczegółowy opis w tekście.

Inicjacja translacji wymaga obecności rozdzielonych podjednostek rybosomów. U organizmów eukariotycznych inicjacja translacji to proces utworzenia gotowego do syntezy białka rybosomu 80S, gdzie kodon inicjujący jest sparowany z inicjatorowym tRNA^{Met} w miejscu P w rybosomie. Ten najbardziej skomplikowany, dwuetapowy proces wymaga przynajmniej dziewięciu eukariotycznych czynników inicjujących eIF (ang. *eukaryotic initiation factor*). Pierwszym etapem inicjacji translacji jest utworzenie kompleksu inicjującego 48S, gdzie kodon inicjujący jest sparowany z inicjatorowym tRNA^{Met} w miejscu P podjednostki 40S rybosomu. Drugi polega na połączeniu kompleksu 48S z podjednostką 60S i uformowaniu gotowego do translacji rybosomu 80S (Jackson et al., 2010; Shirokikh i Preiss, 2018).

W przypadku większości mRNA, kompleks 48S formuje się w wyniku mechanizmu skanowania. Najpierw powstaje kompleks preinicjujący PIC (ang. *preinitiation complex*) o masie 43S (Fig 1b), składający się z podjednostki 40S rybosomu, tymczasowego kompleksu eIF2-TC (eIF2-GTP-tRNA^{Met}) (Fig 1a), eIF3, eIF1, eIF1A oraz eIF5. Do struktury kapu 5' wiążą się eIF4A, eIF4B, eIF4E, eIF4G wraz z białkiem wiążącym ogon poli(A) PABP (ang. *poly(A)-binding protein*) (Fig 1c), a następnie rekrutują PIC (Fig 1d), który rozpoczyna skanowanie 5'UTR mRNA w kierunku 5'→3' (Fig 1e) aż do znalezienia kodonu inicjującego (Fig 1f). Po jego rozpoznaniu formuje się kompleks 48S, a eIF5 i eIF5B aktywują hydrolizę GTP związanego z eIF2, po czym czynniki inicjujące są zastępowane i dołączana jest podjednostka 60S rybosomu (Fig 1g-h) (Jackson et al., 2010; Lind et al., 2017; Shirokikh i Preiss, 2018). Oprócz eIF4E również inne czynniki inicjujące, które wiążą strukturę kapu, mogą rekrutować kompleks PIC. W przypadku mRNA c-Jun specyficzna struktura RNA blokuje oddziaływanie eIF4E z kapem, umożliwiając wiązanie innego czynnika inicjującego, eIF3d (Lee et al., 2016).

Inicjacja translacji jest ściśle regulowana, przede wszystkim na dwóch poziomach. Pierwszy z nich polega na kontroli aktywności czynników translacyjnych głównie poprzez fosforylację. Mechanizm ten jest najlepiej poznany dla eIF2 i eIF4F, których fosforylacja jest kluczowa w odpowiedzi na infekcję wirusową. Drugi poziom dotyczy regulacji ilości i stabilności samego mRNA, np. przez cząsteczki mikro RNA (miRNA) oraz różnorodne białka wiążące RNA, w szczególności PABP wiążące ogon poli(A), które bezpośrednio oddziałują z eIF (Jackson et al., 2010). Obecnie dzięki badaniom strukturalnym możemy dokładniej poznać mechanikę działania kompleksu preinicjującego i samego procesu translacji (Lind et al., 2017).

Po ulokowaniu tRNA^{Met} w miejscu P w rybosomie, inicjacja ulega zakończeniu i rozpoczyna się elongacja translacji (Fig 1i), czyli przyłączanie kolejnych aminokwasów i w efekcie synteza białka. Kolejne aminokwasy są dostarczane przez odpowiedni tRNA do miejsca A w rybosomie. Pośredniczy w tym eukariotyczny czynnik elongujący eEF1A (ang. *eukaryotic elongation factor*), aktywowany wcześniej przez związanie GTP. eEF1A-GTP w kompleksie z aa-tRNA wiąże się z miejscem A rybosomu, gdzie następuje parowanie między kodonem w mRNA a antykodonem tRNA, co prowadzi do hydrolizy GTP. Uwalniany jest eEF1A-GDP, a tRNA pozostaje w miejscu A rybosomu. Między aminokwasami w miejscach A i P spontanicznie dochodzi do stworzenia wiązania peptydowego, w czym pośredniczy eIF5A, związany z miejscem E rybosomu. W tym czasie nowopowstały peptyd jest przesuwany z miejsca A w stronę miejsca P, w czym bierze udział eEF2 oraz centrum peptydylowo-transferowe PTC (ang. *peptidyl transferase center*) w dużej podjednostce rybosomu. aa-tRNA, który wcześniej znajdował się w miejscu P, przesuwa się do miejsca E rybosomu, gdzie tRNA jest uwalniany. Do pustego miejsca A przyłącza się kolejny tRNA przy pomocy eEF1-GTP i cykl się powtarza (Dever et al., 2018).

Ostatnim etapem translacji jest terminacja (Fig 1j) i recykling podjednostek rybosomu i mRNA (Fig 1k), który rozpoczyna się, gdy w miejscu A rybosomu znajdzie się kodon STOP. Za rozpoznanie kodonów STOP odpowiada eukariotyczny czynnik uwalniający eRF1 (ang. *eukaryotic release factor 1*). Energię potrzebną do przeprowadzenia reakcji dostarcza eRF3 wiążący GTP. Kompleks eRF1-eRF3-GTP przyłącza się do kodonu STOP, następuje hydroliza GTP i uwolnienie nowopowstałego peptydu, a kompleks eRF1-eRF3-GDP pozostaje związany do rybosomu, tworząc wraz z nim kompleks post-terminacyjny post-TC (ang. *post-termination complex*). eRF3-GDP oddysocjowuje, a pozostały w post-TC eRF1 bierze udział w recyklingu rybosomu (Bertram et al., 2001; Jackson et al., 2012).

Translacja jest procesem cyklicznym i podjednostki rybosomowe, które w niej uczestniczą podlegają recyklingowi po zakończeniu syntezy białka (Jackson et al., 2010; Jackson et al., 2012). Recyklingowi podlega również mRNA, które służy jako matryca dla kolejnych białek (Gunišová et al., 2018). Recykling potranslacyjny wymaga jonów Mg²⁺ oraz ATPazy ABCE1 (ang. *ATP-binding cassette subfamily E member 1*) wiążącej kasetę ATP z podrodziny E. ABCE1 wraz z eRF1 rozdziela rybosom na wolną podjednostkę 60S, tRNA oraz związaną z mRNA podjednostkę 40S (Jackson et al., 2010).

2. Niekanoniczna translacja

Niekanoniczna translacja jest terminem niezwykle szerokim, obejmującym zarówno mechanizmy niekanonicznej inicjacji, elongacji i terminacji (Fig 2). Proces ten został po raz pierwszy opisany u wirusów, gdzie jest niezwykle ważny dla cyklu replikacyjnego i wirulencji, gdyż jest jednym ze sposobów upakowania i wykorzystania małego genomu (Firth i Brierley, 2012). Coraz więcej badań wskazuje na to, że podobne mechanizmy działają także u innych organizmów. Wykorzystanie niekanonicznej translacji, tak jak alternatywnych miejsc startu transkrypcji oraz alternatywnego splicingu pre-mRNA, pozwala na zwiększenie różnorodności proteomu przy użyciu mniejszej liczby genów (Touriol et al., 2003; Keren et al., 2010; de Klerk i 't Hoen, 2015; Dvinge, 2018). Do tej pory opisano różne mechanizmy niekanonicznej inicjacji, elongacji i terminacji translacji, które mogą się ze sobą łączyć (Fig 2).

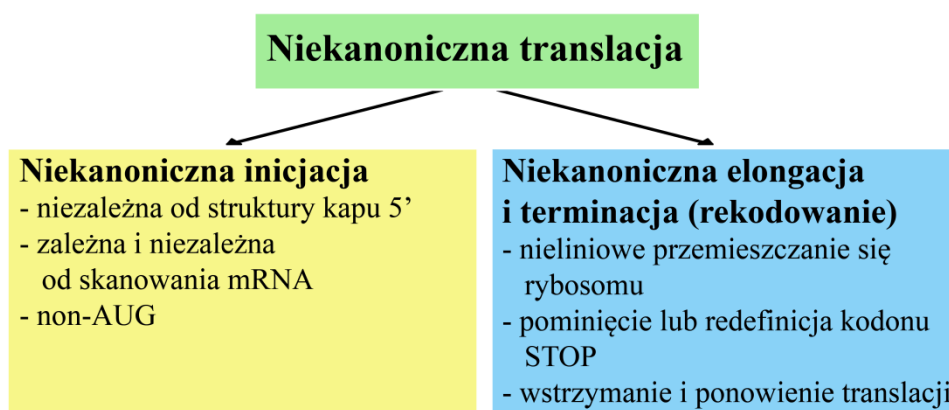


Fig 2. Rodzaje niekanonicznej translacji. Dokładny opis poszczególnych mechanizmów w tekście.

2.1. Niekanoniczna inicjacja translacji

Wśród mechanizmów niekanonicznej inicjacji translacji można wyróżnić niezależne od struktury kapu 5', zależne i niezależne od skanowania mRNA oraz inicjację z niekanonicznych kodonów non-AUG.

2.1.1. Mechanizmy translacji niezależne od struktury kapu 5'

Kluczowe dla inicjacji translacji jest przyłączenie podjednostki 40S rybosomu do struktury kapu 5' mRNA. Oddziaływanie małej podjednostki rybosomu z mRNA może także zachodzić niezależnie od kapu, m.in. za pośrednictwem struktur drugorzędowych, takich jak IRES (ang. *internal ribosome entry site*) czy CITE (ang. *cap-independent translation enhancers*), albo modyfikacji RNA (np. m⁶A), które wiążą czynniki translacyjne i rybosom (Sriram et al., 2018).

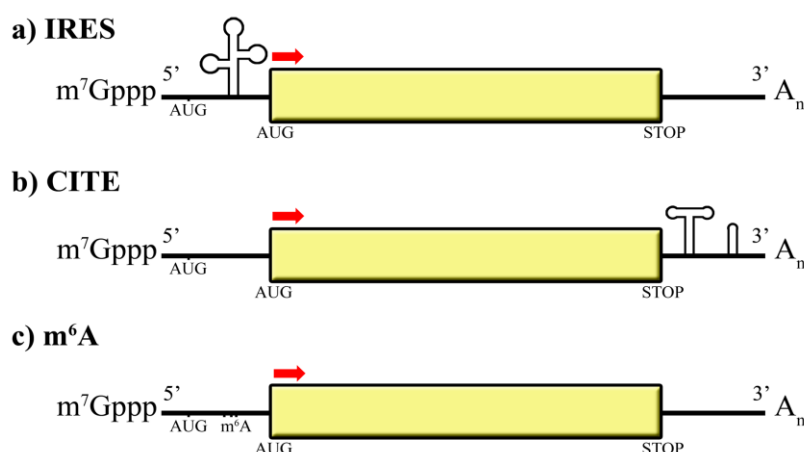


Fig 3. Mechanizmy inicjacji translacji niezależne od struktury kapu. Czerwoną strzałką zaznaczono inicjację translacji, żółty prostokąt oznacza ORF (ang. *open reading frame*) ulegający translacji. Szczegółowy opis w tekście.

Wewnętrzne miejsca wiązania rybosomu IRES (Fig 3a) to umiejscowione w 5'UTR mRNA specyficzne struktury RNA, które są rozpoznawane przez rybosomy. Zostały po raz pierwszy odkryte u *Picornaviridae* i obecnie wyróżnia się ich cztery klasy (Mailliot i Martin, 2018; Martinez-Salas et al., 2018). Na poziomie sekwencji wirusowe IRES nie są szczególnie konserwowane ewolucyjnie, ale na poziomie struktury wykazują znaczące podobieństwo (Martinez-Salas et al., 2018). Są powszechne w mRNA wirusów atakujących komórki eukariotyczne, ponieważ imitują kap 5' do którego wiążą się rybosomy (Firth i Brierley, 2012). W ten sposób pozwalają obejść komórkowe zabezpieczenia inicjacji translacji zależnej od kapu, prowadząc do produkcji białek wirusowych i replikacji wirusa (Sriram et al., 2018). Z powodu swojej charakterystycznej budowy, wirusowe IRES są brane pod uwagę jako cele w terapiach przeciwwirusowych. Obecnie bada się działanie modyfikowanych antysensownych oligonukleotydów, siRNA, morfolinowych oligomerów, aptamerów i małych cząsteczek, będących pochodnymi benzimidazolu (Martinez-Salas et al., 2018).

IRES występują też naturalnie u organizmów komórkowych, zarówno bakteryjnych jak i jądrowych (Touriol et al., 2003; Kearse i Wilusz, 2017). Około 10% genów w komórkach ludzkich posiada elementy IRES, z których większość znajduje się w rejonach 5'UTR, lecz zaobserwowano je również w 3'UTR (Weingarten-Gabbay et al., 2016; Kwan i Thompson, 2018). Zostały wykryte głównie w mRNA czynników transkrypcyjnych (21%), czynników wzrostu (15%) oraz transporterów, receptorów i kanałów błonowych (22%) (Lacerda et al., 2017). IRES rekrutują zarówno kanoniczne jak i niekanoniczne czynniki translacyjne, co w efekcie prowadzi do związania rybosomu. W komórkach eukariotycznych specjalną klasą białek są czynniki ITAF (ang. *IRES trans-acting factors*) oddziałujące z IRES

in trans, które poprzez oddziaływanie z IRES przyczyniają się do rekrutacji rybosomu i stabilizacji 40S (Lacerda et al., 2017; Sriram et al., 2018).

Eukariotyczne IRES dzielą się na dwie grupy. IRES typu „przyłącz i skanuj” (ang. “*land and scan*”) znajdują się niedaleko kodonu start i rekrutują podjednostkę 40S, która następnie skanuje mRNA do najbliższego kodonu AUG. Do tej grupy zalicza się m.in. IRES białek z rodziny Myc (Spriggs et al., 2009). Druga klasa IRES zawiera sekwencje podobne do bakteryjnych sekwencji Shine-Dalgarno, które pomagają rekrutować rybosom do mRNA, np. w mRNA *IGF1R* (ang. *insulin-like growth factor 1 receptor*) (Meng et al., 2010). Zadaniem elementów IRES w komórce jest przede wszystkim umożliwienie adaptacji do warunków stresowych, podczas których translacja zależna od kapu jest hamowana (Mitchell i Parker, 2015). Przeważnie mRNA komórkowe zawierające IRES podlegają translacji w kanoniczny sposób, a ich IRES używane są w warunkach stresu (Kwan i Thompson, 2018). Dodatkowo IRES wspomagają translację mRNA o silnie ustrukturalizowanych UTR, do których kompleks wiążący kap ma utrudniony dostęp (Sriram et al., 2018).

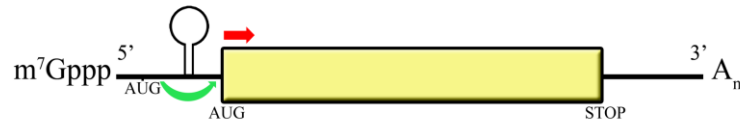
Kolejnym mechanizmem inicjacji syntezy białka, w którym nie pośredniczy kap, jest wykorzystanie niezależnych od kapu wzmacniaczy translacji CITE (Fig 3b), które zostały opisane po raz pierwszy w genomach roślinnych wirusów z rodziny *Tombusviridae* (Wu i White, 1999). CITE znajdują się w 3'UTR wirusowych mRNA i mogą przybierać rozmaite struktury w kształcie liter T, Y, I lub pseudowęzła, w celu rekrutacji kompleksu eIF4F i podjednostek rybosomu. W przeciwieństwie do IRES wymagają wolnych końców 5' mRNA, które ściśle regulują ich działanie (Sriram et al., 2018; Kwan i Thompson, 2018). CITE są trudne do wykrycia z powodu swojej zmiennej struktury. Zwykle bada się je poprzez przyłączenie do genu reporterowego, na przykład lucyferazy. Obecnie wyróżnia się siedem klas CITE, różniących się strukturą i charakterystycznych dla danych rodzin wirusów (Truniger et al., 2017). CITE są samodzielnymi elementami, które mogą być przenoszone między wirusami w wyniku rekombinacji homologicznej, stanowiąc ważny element adaptacji wirusów i ich ewolucji (Truniger et al., 2017). W przypadku wirusów roślinnych CITE zwiększają stopień ich wirulencji poprzez wspomaganie wydajności namnażania białek wirusowych (Du et al., 2017). Elementy CITE obserwuje się również w komórkach zwierzęcych (Shatsky et al., 2018), gdzie przeważnie występują na końcach 3' mRNA, ale są też obecne na końcach 5' (Lacerda et al., 2017).

W przypadku inicjacji translacji eukariotycznych mRNA, niezwykle istotne są również modyfikacje RNA, a zwłaszcza N⁶-metylacja adenozyiny (m⁶A) (Fig 3c). Jest to najpowszechniejsza (ponad 80%) z podstawowych modyfikacji RNA. Początkowo m⁶A wykrywano głównie w 3'UTR mRNA w okolicach kodonów STOP (93% wszystkich m⁶A w UTR) (Meyer et al., 2012), ale w ostatnich latach zidentyfikowano je również w 5'UTR i stwierdzono, że mają wpływ na inicjację translacji. Wykazano, że rozpoznanie m⁶A w tym regionie przez białko YTHDF1 (ang. *YTH domain-containing family protein 1*) stymuluje wiązanie czynnika eIF3 do m⁶A i w rezultacie promuje translację. Natomiast deplecja zarówno metylazy METTL3 odpowiedzialnej za syntezę tych modyfikacji, jak i YTHDF1, selektywnie inhibuje translację (Coots et al., 2017; Sriram et al., 2018). Miejsca do których wiąże się rybosom dzięki modyfikacjom m⁶A nazwano indukowanymi przez m⁶A miejscami przyłączania rybosomów MIREs (ang. *m⁶A-induced ribosome engagement sites*) (Lacerda et al., 2017). W istocie, mRNA pozbawione struktury kapu 5' a zawierające m⁶A mogą ulegać translacji w systemach pozakomórkowych *in vitro*. Co więcej, ponieważ m⁶A bezpośrednio wiąże eIF3, inicjacja translacji zachodzi pod nieobecność kompleksu eIF4F. Oznacza to, że modyfikacje m⁶A, podobnie jak IRES, mogą promować syntezę białka nawet w warunkach stresowych (Mitchell i Parker, 2015; Shirokikh i Preiss, 2018). Interesujące jest, że modyfikacje m⁶A wraz z IRES przyczyniają się do translacji kolistych cząsteczek RNA (circRNA, ang. *circular RNA*) w komórkach ludzkich. W przypadku m⁶A działa podobny mechanizm polegający na rekrutacji kompleksu preinicjującego przez białko YTHDF1 (Yang et al., 2017).

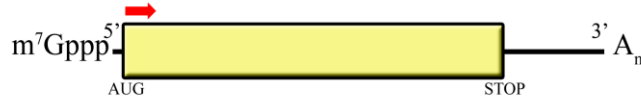
2.1.2. Mechanizmy inicjacji zależne i niezależne od skanowania mRNA

Inicjacja translacji może zachodzić w sposób niekanoniczny także w kolejnych etapach, po związaniu PIC ze strukturą kapu (Fig 4). Jednym z mechanizmów niezależnych od tradycyjnego skanowania jest przełączanie rybosomów (ang. *ribosome shunting*), kiedy rybosomy przemieszczają się nieliniowo po mRNA, pomijając niektóre obszary (Fig 4a). Pierwszy raz zjawisko to zaobserwowano u wirusa mozaiki kalafiora CaMV (ang. *Cauliflower mosaic virus*) w sekwencji liderowej pgRNA (ang. *pregenomic RNA*), która przyjmuje strukturę drugorzędową spinki do włosów, pomijaną przez rybosomy. Region poniżej tej struktury jest bogaty w pary UA, co stanowi miejsce docelowe dla przełączonego rybosomu (Fütterer et al., 1990; Pooggin i Ryabova, 2018).

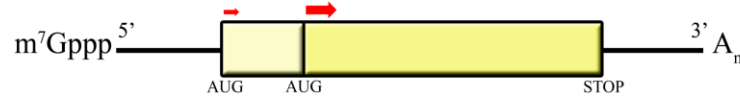
a) przełączanie rybosomów



b) TISU



c) przepuszczające skanowanie



d) reinicjacja

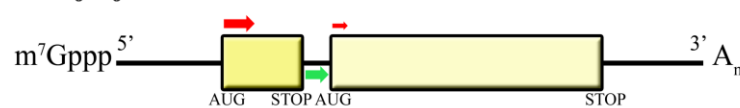


Fig 4. Mechanizmy inicjacji niezależne i zależne od skanowania mRNA. Czerwoną strzałką zaznaczono inicjację translacji, grubość strzałki odpowiada sile translacji. Żółty prostokąt oznacza ORF ulegający translacji, różne odcienie pokazują sąsiadujące ORFy. Zielone strzałki opisują przemieszczanie się rybosomu w sposób niekanoniczny. Szczegółowy opis w tekście.

O ile wykorzystanie IRES pozwala na dołączenie rybosomu niezależnie od kapu, to przełączanie rybosomów uniezależnia je od kanonicznego mechanizmu skanowania (Firth i Brierley, 2012) (Fig 5). Przełączanie rybosomów obserwuje się również u organizmów komórkowych (Sriram et al., 2018). Wiele mRNA dla których zachodzi przełączanie rybosomów koduje nie tylko standardowy ORF, ale posiada także jedno lub więcej krótkich ORF (ang. *short ORF*; sORF). Po terminacji translacji sORF, rybosom pozostaje związany z mRNA i może pominąć kolejne sORF czy struktury drugorzędowe i wznowić translację na położonym poniżej głównym ORF (ang. *major ORF*, mORF) (Kwan i Thompson, 2018).

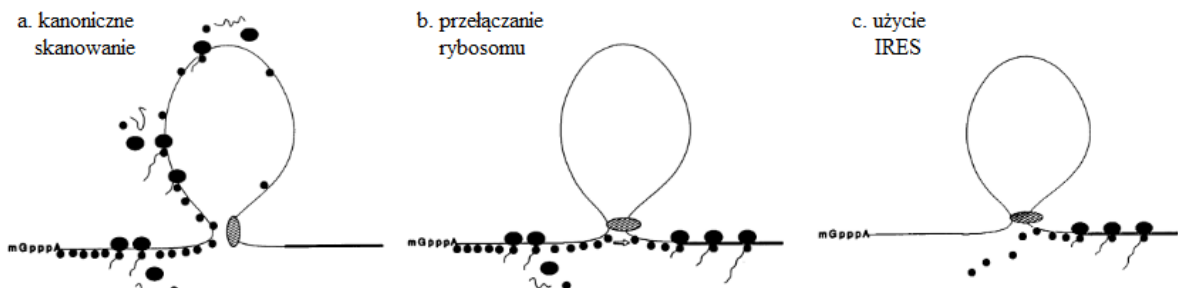


Fig 5. Porównanie mechanizmów niekanonicznej inicjacji translacji, opartych na strukturze drugorzędowej RNA. a) kanoniczne skanowanie; b) przełączanie rybosomów; c) IRES. Mniejsze kropki pokazują małą podjednostkę rybosomu podczas skanowania mRNA; większe kropki obrazują dużą podjednostkę rybosomu, której przyłączenie do PIC rozpoczyna translację. Na szaro zaznaczono białka wiążące i stabilizujące strukturę spinki do włosów. Na podstawie (Fütterer et al., 1990).

Zarówno w przypadku IRES jak i przełączania rybosomów wymagana jest obecność skomplikowanych struktur drugorzędowych, stąd występują one raczej w mRNA o długich sekwencjach 5'UTR. Jednak u wyższych eukariontów również mRNA o bardzo krótkich 5'UTR mogą ulegać wydajnej translacji, podczas której za inicjację odpowiadają elementy TISU (ang. *translation initiator of short 5' UTRs*), a proces ten zachodzi bez skanowania. TISU są krótkimi sekwencjami (SAASAUGGCGGC), które znajdują się w regionie od +5 do +30 od miejsca startu transkrypcji TSS (ang. *transcription start site*). TISU zostały wykryte dzięki badaniom bioinformatycznym rejonów promotorowych ludzkich mRNA. Obecne są przede wszystkim w genach o fundamentalnych funkcjach w komórce, takich jak metabolizm białek. Obecność TISU pozwala na redukcję 5'UTR do minimum, nawet do 5 nukleotydów (Elfakess i Dikstein, 2008; Haimov et al., 2015). Ciekawe jest, że w przypadku inicjacji translacji za pomocą TISU wymagane jest działanie eIF1, który normalnie blokuje start syntezy białka z kodonów położonych zbyt blisko struktury kapu (Hinnebusch et al., 2016).

Elementy podobne do TISU wykryto również u niższych eukariontów, takich jak *Giardia lamblia*, *Entamoeba histolytica*, *Trichomonas vaginalis*, czy *Saccharomyces cerevisiae*. Szczególnie ciekawy przypadek stanowi tu *G. lamblia*, którego mRNA posiadają pozbawione sekwencji konsensusowych bardzo krótkie 5'UTR (od 0 do 14 nukleotydów). Jednak nawet tak krótkie sekwencje na końcu 5' mRNA są kluczowe dla inicjacji, ponieważ ich wydłużenie prowadzi do znaczącego spadku wydajności translacji (Li i Wang, 2004; Haimov et al., 2015), podczas gdy u innych eukariontów efektywność translacji przeważnie wzrasta wraz z długością 5'UTR.

Nie tylko obecność struktury drugorzędowej w sekwencji mRNA prowadzi do pominięcia niektórych kodonów startu. Czasami kompleks preinicjujący przepuszcza niektóre kodony, gdyż nie znajdują się w odpowiednim kontekście nukleotydowym Kozak (Kozak, 1987), który u ssaków stanowi sekwencja GCCRCCAUGG (gdzie R oznacza purynę). Mechanizm ten, zwany „przepuszczającym” czy „nieszczelnym” skanowaniem (ang. *leaky scanning*) sprawia, że rybosom nie rozpoczyna translacji od kodonu startu poprzedzonego słabą sekwencją Kozak, ale kontynuuje skanowanie w poszukiwaniu kolejnego kodonu o silniejszym kontekście (Fig 4c) (Kozak, 2002).

Przepuszczające skanowanie jest jednym z najbardziej powszechnych niekanonicznych mechanizmów inicjacji translacji dla wirusowych mRNA i może prowadzić do powstania kilku izoform tego samego białka, ale także reguluje syntezę białek

u organizmów eukariotycznych (Firth i Brierley, 2012). Jeżeli położony powyżej dodatkowy kodon AUG jest częścią uORF (ang. *upstream ORF*), to mechanizm ten wpływa na translację mORF położonego poniżej. Dwa kodony startu w tej samej ramce odczytu pozwalają na syntezę dwóch izoform z różną wydajnością, w zależności m.in. od kontekstu Kozak. Natomiast w przypadku kodonów inicjujących translację w różnych ramkach odczytu, z tej samej cząsteczki mRNA powstają dwa różne białka (Sriram et al., 2018).

Po dotarciu przez rybosom do kodonu STOP następuje terminacja translacji, polegająca na uwonieniu powstałego białka i oddysocjowanie obu podjednostek rybosomu od mRNA (Bertram et al., 2001; Jackson et al., 2012). Zdarza się jednak, zwłaszcza po zakończeniu syntezy krótkich peptydów (mniej niż 30 aminokwasów), że mała podjednostka 40S pozostaje przyłączona do mRNA i kontynuuje skanowanie w poszukiwaniu kolejnego miejsca inicjacji. Reinicjacja translacji (ang. *reinitiation*), czyli ponowna inicjacja translacji (Fig 4d), zachodzi często, zwłaszcza u ssaków, dla mRNA w których obecne są uORF i wymaga obecności czynników inicjujących związanych z rybosomem podczas translacji uORF. Ponadto, ORF podlegający reinicjacji musi znajdować się pewnej odległości od poprzedniego, gdyż mała podjednostka rybosomu nie może rozpocząć kolejnej inicjacji zaraz po terminacji poprzedniej ramki odczytu. Kluczowa jest również długość pierwszego produktu, ponieważ reinicjacja po translacji długiego ORF (np. kodującego białko) jest zdecydowanie rzadsza i wymaga specyficznych sygnałów w mRNA lub specyficznych czynników działających *in trans* (Jackson et al., 2012; Firth i Brierley, 2012; Gunišová et al., 2018). W komórkach ludzkich, gdzie aż 49% genów posiada uORF, w reinicjacji biorą udział kanoniczny czynnik translacyjny eIF3h oraz niekanoniczne czynniki eIF2D, DENR i MCTS1. Istotnym elementem tego procesu jest ponowna rekrutacja inicjującego tRNA^{Met} do małej podjednostki rybosomu w sposób niezależny od eIF2 (Sriram et al., 2018).

U niektórych wirusów, takich jak koci kaliciwirus FCV (*feline calicivirus*), występują elementy TURBS (ang. *termination upstream ribosome binding site*). Są to długie ~90 nukleotydowe sekwencje, które stabilizują rybosom podczas procesu reinicjacji. Elementy TURBS posiadają trzy motywy, które odpowiadają za wiązanie 18S rRNA, uniemożliwiają oddysocjowanie rybosomu, stabilizują strukturę drugorzędową oraz pozycjonują rybosom na kodonie start nowego ORF (Luttermann i Meyers, 2014; Pooggin i Ryabova, 2018). Reinicjacja poprzez TURBS zmniejsza znaczenie eIF3 w recyklingu podjednostek 40S rybosomu, umożliwiając udział eIF2/1/1A lub eIF2D po zależnym lub niezależnym od ATPazy ABCE1 rozłożeniu i recyklingu kompleksu post-terminacyjnego.

W reinicjacji mogą zostać również użyte kodony niekanoniczne, ale preferowany jest standardowy AUG (Zinoviev et al., 2015).

2.1.3. Inicjacja translacji z niekanonicznych kodonów non-AUG

Kolejnym przykładem nietypowej inicjacji translacji jest wykorzystanie niekanonicznych kodonów startu nazywanych kodonami non-AUG. Teoretycznie każdy z kodonów w genomie może zostać użyty jako alternatywny kodon inicjujący, jednak największą wydajność inicjacji wykazują kodony spokrewnione (ang. *near-cognate codons*), czyli różniące się od AUG jedną zasadą azotową (Chang et al., 2010; Hecht et al., 2017). Przeważnie inicjacja translacji z kodonów niekanonicznych zachodzi na znacznie niższym poziomie niż ze standardowego kodonu AUG (Michel et al., 2014).

Inicjacja translacji z kodonu innego niż AUG została po raz pierwszy zaobserwowana u wirusa Sendai (Curran i Kolakofsky, 1988) oraz wirusa mysiej białaczki Mo-MuLV (Prats et al., 1989). W obu przypadkach niekanoniczne starty translacji znajdują się powyżej kodonu AUG, z tym, że dla wirusa Sendai kodon ACG jest położony w innej ramce odczytu, a dla wirusa Mo-MuLV kodon CUG w tej samej. Te dwa odkrycia ujawniły istnienie nowych możliwości regulacji ekspresji genów, zarówno pod kątem mechanizmu działania jak i biologicznego znaczenia startu translacji z kodonów non-AUG. Do niedawna znalezienie takich startów było dziełem przypadku, ale obecnie dzięki profilowaniu rybosomów (ang. *ribosome profiling*), możliwe stało się określenie powszechności tego zjawiska (Kearse i Wilusz, 2017).

Podczas inicjacji translacji z niekanonicznych kodonów non-AUG każdy z wcześniej wymienionych mechanizmów może prowadzić do wyboru kodonu non-AUG na miejsce startu syntezy białka (Fig 6). Istnieje wiele przypadków, w których za wykorzystanie niekanonicznego kodonu w inicjacji translacji odpowiadają wewnętrzne miejsca IRES (Fig 6a). Niekiedy wiązanie rybosomu do IRES jest wspomagane przez czynniki translacyjne, takie jak DAP5 (także zwane jako eIF4G2 albo NAT1), które ogrywają kluczową rolę podczas zależnego od IRES startu translacji z kodonu GUG w komórkach mysich i ludzkich (Lieberman et al., 2015). IRES zlokalizowany tuż przed kodonami non-AUG jest obecny w 5'UTR mRNA kodujących ludzkie czynniki wzrostu FGF2 i VEGF. Podobnie jest w przypadku wielu genów wirusowych, gdzie elementy IRES bezpośrednio wpływają na wybór niekanonicznych kodonów na miejsca startu syntezy białka (Touriol et al., 2003).

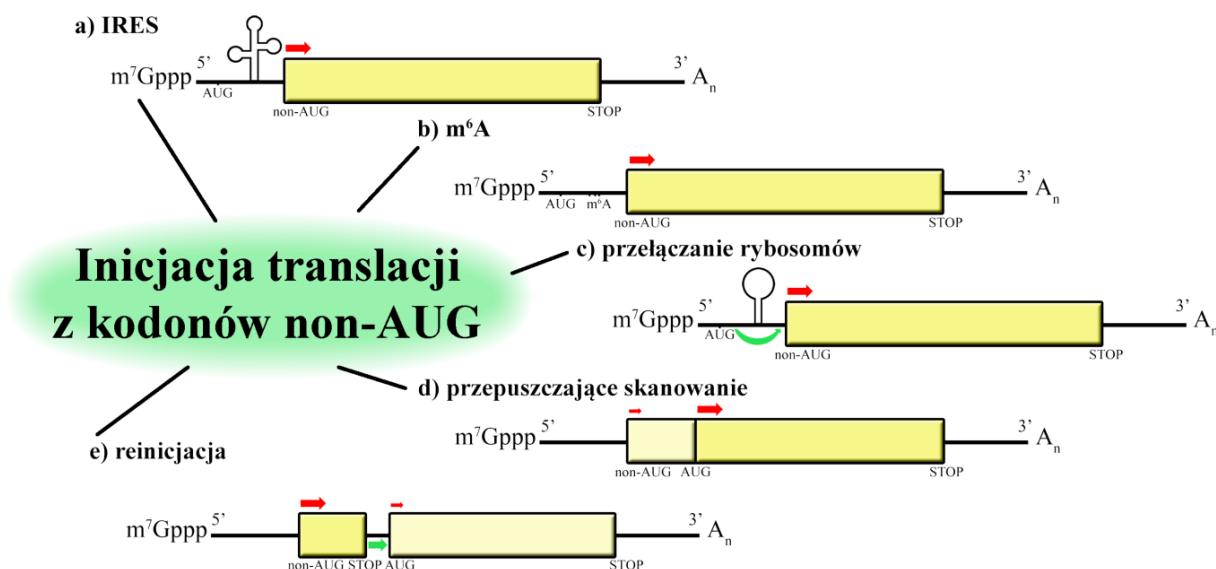


Fig 6. Niekanoniczne mechanizmy translacji, odpowiadające za wybór kodonów non-AUG na miejsca startu syntezy białka. Czerwoną strzałką zaznaczono inicjację translacji, grubość strzałki odpowiada sile translacji. Żółty prostokąt oznacza ORF ulegający translacji, różne odcienie pokazują sąsiadujące ORFy. Zielone strzałki opisują przemieszczanie się rybosomu w sposób niekanoniczny. Szczegółowy opis w tekście.

Na wybór kodonów non-AUG, zwłaszcza w warunkach stresowych, mają również wpływ modyfikacje m^6A w obrębie mRNA (Fig 6b). Globalne profilowanie m^6A i miejsc inicjacji translacji podczas głodzenia pokazało zależność między m^6A obecnymi w 5'UTR a wzrostem inicjacji translacji z kodonów niekanonicznych (Zhou et al., 2018; Meyer, 2018). Start translacji z kodonów niekanonicznych może być także promowany przez odpowiedni kontekst nukleotydowy, który jest zbliżony do klasycznego kontekstu Kozak. Każdy z kodonów non-AUG ma swój preferowany kontekst i położone w innym będą wykorzystywane ze znacznie niższą wydajnością lub zupełnie pomijane (Chen et al., 2008).

Obecnie uważa się, że jednym z głównych mechanizmów wyboru kodonów non-AUG na miejsce startu translacji jest mechanizm przepuszczającego skanowania (Fig 6d). Dla kodonów niekanonicznych kontekst nukleotydowy jest znacznie bardziej istotny niż w przypadku kodonu AUG (Michel et al., 2014; Hinnebusch et al., 2016). Wybór kodonów niekanonicznych w mechanizmie przepuszczającego skanowania w dużym stopniu jest dyktowany stosunkiem między czynnikami inicjującymi eIF1 i eIF5. Zadaniem eIF1 jest promowanie skanowania i pomijanie kodonów inicjujących, zarówno non-AUG jak i AUG, w słabym kontekście Kozak, zaś eIF5 działa antagonistycznie. Zależność między eIF1 i eIF5 jest wykorzystywana w procesach autoregulacyjnych w komórce i w odpowiedzi na stres (Hinnebusch et al., 2016).

Na wybór kodonu inicjującego przez PIC mają również wpływ obecne poniżej struktury drugorzędowe RNA (Fig 6c). Na przykład inicjacja z kodonu non-AUG lub kodonu AUG w słabym kontekście Kozak może być promowana dzięki obecności termostabilnej struktury spinki do włosów w mRNA, położonej 14 nukleotydów poniżej kodonu inicjującego. Struktura ta powoduje zatrzymanie rybosomu, w którego miejscu P umieszczony zostaje kodon inicjujący, co sprzyja jego wykorzystaniu zarówno *in vitro* jak i *in vivo* (Kozak, 1990; Liang et al., 2017).

Podczas stresu syntetyzowanych jest wiele uORF, które mogą wpływać na translację głównej ramki odczytu (Fig 6e) i których synteza jest często inicjowana z niekanonicznych kodonów startu. Tak jest na przykład w przypadku drożdżowego genu *GCN4*, który koduje aktywator transkrypcji bZIP, indukujący ekspresję większości enzymów biosyntezy aminokwasów w odpowiedzi na głodzenie aminokwasowe. Synteza białka Gcn4 jest silnie regulowana zarówno na poziomie białka jak i mRNA (Hinnebusch, 2005). W mRNA *GCN4* zawarte są liczne uORF, również takie rozpoczynające się od kodonu non-AUG. W zależności od warunków stresowych następuje synteza poszczególnych uORF, które zależnie od poziomu fosforylacji eIF2 α , wpływają na translację mORF (Holcik i Sonenberg, 2005, Zhang i Hinnebusch, 2011).

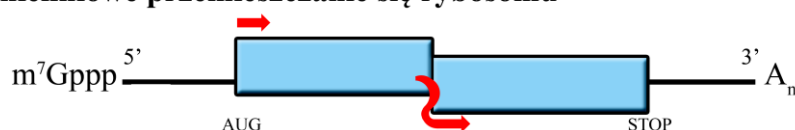
Obecnie nie są znane przypadki elementów CITE odpowiadających za wybór niekanonicznego kodonu inicjacji translacji. Co więcej, w badaniach nad wirusowym CITE pokazano, że zmiana kodonu AUG na GUG skutkuje inhibicją translacji w wyniku inicjacji translacji zależnej od 3'CITE (Miras et al., 2018). Podobnie jest z elementami TISU, które w swojej sekwencji posiadają kodon AUG i docelowo służą do syntezy białek, które są produkowane na wysokim poziomie (Haimov et al., 2015).

Ciekawe są też sytuacje, kiedy jedynym produktem translacji jest białko inicjowane z kodonu non-AUG. Analiza dostępnych danych z profilowania rybosomów u drożdży *Saccharomyces cerevisiae* pozwoliła na wykrycie w standardowych warunkach 468 przypadków, gdzie inicjacja translacji głównej formy białka najprawdopodobniej rozpoczyna się z kodonu spokrewnionego z AUG (Monteuuis et al., 2019). Nie przeprowadzono jednak odpowiednich analiz, które mogłyby ujawnić obecność optymalnego kontekstu nukleotydowego, struktur IRES czy innych elementów promujących wybór kodonów niekanonicznych.

2.2. Niekanoniczna elongacja i terminacja translacji (rekodowanie)

Niekanoniczne mechanizmy translacji mające miejsce podczas elongacji lub terminacji translacji nazywa się również mechanizmami rekodowania (ang. *recoding*) (Baranov et al., 2006) (Fig 7). Zalicza się do nich procesy polegające na nieliniowym przemieszczaniu się rybosomu podczas elongacji. Pierwszym z nich jest programowana zmiana ramki odczytu PRF (ang. *programmed ribosomal frameshifting*), która jest wynikiem przesunięcia się rybosomu o jeden lub dwa nukleotydy i wznowieniu translacji w innej ramce odczytu. Drugim jest pominięcie przez rybosom większej części transkryptu i kontynuacja translacji (ang. *ribosomal bypassing*). Kolejnym z mechanizmów niekanonicznej elongacji i terminacji jest pominięcie kodonu STOP (ang. *stop-codon readthrough*), prowadzące do syntezy białka wydłużonego na końcu C. Niekiedy zamiast kodonu STOP mogą zostać inkorporowane niestandardowe aminokwasy, takie jak selenocysteina czy pirolizyna, co nazywane jest zmianą definicji kodonu STOP (ang. *stop-codon redefinition*). Ostatnim mechanizmem jest wstrzymanie i ponowienie translacji (ang. „*Stop-Carry On*” lub „*StopGo*” translation), prowadzące do powstania kilku białek z jednego ORF.

a) nieliniowe przemieszczanie się rybosomu



b) pominięcie lub redefinicja kodonu STOP



c) wstrzymanie i ponowienie translacji



Fig 7. Mechanizmy niekanonicznej elongacji i terminacji translacji. Czerwoną strzałką oznaczono start translacji oraz niekanoniczne przemieszczanie się rybosomu. Niebieski prostokąt wskazuje na ORF ulegający translacji. Znaki (X) i (▶) oznaczają odpowiednio miejsca wstrzymania i ponowienia translacji. Szczegółowy opis w tekście.

Rekodowanie jest niezwykle powszechne w genach prokariotycznych o wirusowym i mobilnym pochodzeniu (Baranov et al., 2006). Większość mechanizmów rekodowania została opisanych w bazie RECODE (Bekaert et al., 2010), która jednak nie uwzględnia inkorporacji pirolizyny oraz zjawiska wstrzymania i ponowienia translacji.

Programowana zmiana ramki odczytu w trakcie translacji (PRF) (Fig 7a) polega na zmianie ramki odczytu podczas elongacji translacji na ramkę -1, -2 lub +1 (Baranov et al., 2006). PRF po raz pierwszy zaobserwowano u wirusa mięsaka Rousa (*Rous sarcoma alpharetrovirus*), który w swoim genomie posiada dwa nakładające się ORF: *gag* i *pol*, ulegające translacji jako poliproteina Gag-Pol (Jacks i Varmus, 1985; Firth i Brierley, 2012). Podobny mechanizm obserwuje się też i u innych wirusów, łącznie z ludzkim wirusem niedoboru odporności (*human immunodeficiency virus*) HIV-1 i HIV-2. Sygnałem PRF jest sekwencja X_XXY_YYZ (gdzie podkreślenia separują ramkę odczytu 0; X oznacza trzy takie same nukleotydy, Y to A lub U, a Z reprezentuje A, C lub U), na której rybosom zatrzymuje się i zmienia ramkę odczytu. Po „prześlizgnięciu się” rybosomu o jeden nukleotyd wstecz, powstaje ramka odczytu XXX_YYY_Z, z której kontynuowana jest translacja (Giedroc i Cornish, 2009).

PRF zachodzi również u organizmów komórkowych. W genomie drożdży *S. cerevisiae* retrotranspozon *Ty* jest kodowany przez geny *TYA* i *TYB*. W genie *TYB*, zawierającym sekwencję homologiczną do retrowirusowego genu *pol*, znajduje się sygnał PRF CUU_AGG_C, gdzie następuje zmiana odczytu na ramkę +1 (Belcourt i Farabaugh, 1990). Podobnie zachowują się inne transpozony z rodziny *Ty* (Dever et al., 2018). Spontaniczne zmiany ramki odczytu zachodzą również w drożdżowym mitochondrialnym genie *COX2* (Fox i Weiss-Brummer, 1980).

Oprócz specjalnego motywu kinetyczną pułapkę (ang. *kinetic trap*) do zmiany ramki odczytu może stanowić również kodon STOP. Kluczowa dla PRF jest pauza w translacji, a kodony STOP są naturalnym miejscem zatrzymania rybosomu. U pierwotniaków i w mitochondriach utrata genów odpowiedzialnych za syntezę translacyjnych czynników uwalniających specyficznych dla danego kodonu STOP (najczęściej UGA), prowadzi do odczytania tego kodonu przez supresorowy tRNA, który promuje zmianę ramki odczytu (Lobanov et al., 2017; Dever et al., 2018). PRF mogą wywołać również elementy działające *in cis*, takie jak motyw pseudowęzła (Dever et al., 2018).

Kolejnym niekanonicznym mechanizmem elongacji jest pomijanie długich fragmentów transkryptu przez rybosom (Fig 7a). Mechanizm ten nie zachodzi u wirusów atakujących komórki eukariotyczne (Firth i Brierley, 2012), lecz po raz pierwszy został wykryty w genie 60 bakteriofaga T4 kodującym topoizomerazę (Weiss et al., 1990; Chen et al., 2015). Obecnie wiadomo, że jest to zjawisko powszechne w mitochondriach. Zostało najlepiej zbadane u patogennych grzybów *Magnusiomyces capitatus*, gdzie w genomie

mitochondrialnym wykryto więcej niż 80 pomijanych sekwencji o długości 27-55 nt, które mimo że są obecne w dojrzałym mRNA nie zostały użyte do syntezy białka (Lang et al., 2014). W mechanizmie pomijania fragmentów transkryptu ruch rybosomu nazywany jest przeskakiwaniem (ang. *jumping* lub *hopping*), w odróżnieniu od przełączania rybosomów (ang. *ribosome shunt*) (Nosek, et al., 2015).

Niekiedy białka wydłużone na końcu C powstają w wyniku pominięcia kodonu STOP (Fig 7b). Dochodzi wówczas do rozpoznania kodonu STOP przez spokrewniony lub supresorowy tRNA, po czym translacja jest kontynuowana do kolejnego kodonu STOP. Wyróżnia się trzy motywy promujące pomijanie kodonu STOP, wszystkie z nich są powszechne u różnych typów wirusów (Firth i Brierley, 2012). Terminacja translacji jest procesem bardzo wydajnym i sytuacje pomijania kodonu STOP są niezwykle rzadkie. Co więcej, w przypadku niektórych wirusów oraz genów w komórkach prokariotycznych i eukariotycznych pominięcie kodonu STOP jest integralnym elementem kontroli ekspresji genów. Zjawisko to zależy od kontekstu, w którym znajduje się dany kodon terminacyjny (np. UAG), który może zwiększyć częstość pomijania kodonu STOP z 0,3% do 5% (Bertram et al., 2001; Tate et al., 2018). Znaczenie mają również elementy i motywy obecne poniżej lub powyżej pomijanego kodonu (Firth i Brierley, 2012). Niekiedy wywołanie pominięcia kodonu STOP jest procesem pożądanym. W wielu chorobach genetycznych, takich jak np. dystrofia mięśniowa Duchenne’a, następują spontaniczne mutacje tworzące przedwczesne kodony STOP. W tym przypadku terapia ukierunkowana jest na pomijanie takich kodonów terminacyjnych i syntezę pełnych form białek (Friesen et al., 2018; Dabrowski et al., 2018).

W odróżnieniu od inkorporacji standardowego aminokwasu przy pomijaniu kodonu STOP (ang. *readthrough*), w przypadku zmiany definicji kodonu STOP (ang. *redefinition*) do łańcucha polipeptydowego włączane są niestandardowe aminokwasy (Fig 8), selenocysteina (Sec) oraz pirolizyna (Pyr), nazywane niekiedy 21 i 22 aminokwasem (Longtin, 2004). Zmiana definicji kodonu STOP, podobnie jak pomijanie fragmentów transkryptu przez rybosom, nie występuje u wirusów atakujących komórki eukariotyczne (Firth i Brierley, 2012), ale jest powszechne w komórkach zwierzęcych i ludzkich (Labunsky et al., 2014) a także u bakterii (Baranov et al., 2006) i archeonów (Rother i Krzycki, 2010).

aminokwasowy 2A o sekwencji D₁₂(V/I)ExNPGP₁₉ (Sharma et al., 2012) oraz mniej konserwowane, ale funkcjonalnie ważne aminokwasy znajdujące się powyżej wyjścia z tunelu rybosomu (ogółem około 30 aminokwasów). Przypuszcza się, że struktura poprzedniego peptydu zapobiega wiązaniu się w miejscu A rybosomu tRNA^{Pro}, umożliwiając związanie się eRF1 i eRF3, które uwalniają poprzedni peptyd. Następnie tRNA^{Pro} wiąże się do rybosomu i translacja jest kontynuowana z blisko stuprocentową wydajnością (Brown i Ryan, 2010).

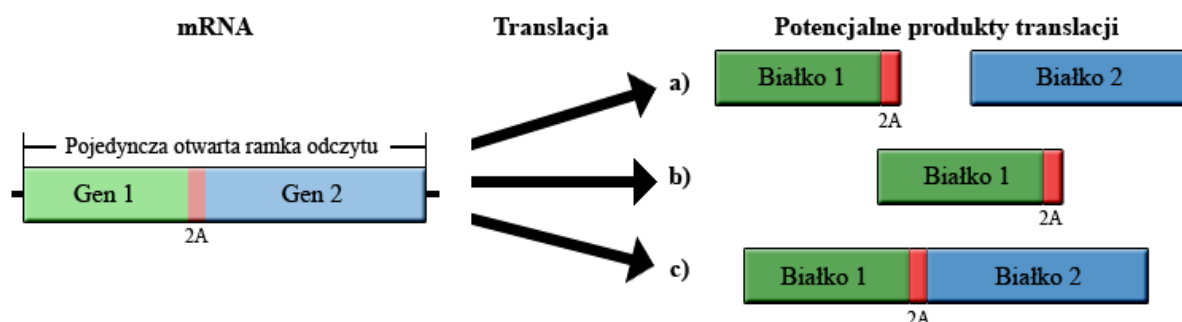


Fig 9. Potencjalne produkty translacji zależnej od motywu 2A (mechanizm wstrzymania i ponowienia translacji). Z jednej otwartej ramki odczytu mogą powstać trzy rodzaje produktu: (a) dwa różne białka, oddzielone od siebie motywem 2A; (b) jedynie pierwsze białko; po wstrzymaniu translacja nie jest kontynuowana; (c) jeden produkt białkowy; wstrzymanie i ponowienie nie następuje. Na podstawie (Luke et al., 2013).

Mechanizm wstrzymania i ponowienie translacji stanowi alternatywę dla proteolitycznego cięcia w produkcji wielu białek z jednego ORF. Takie kasety są obecne w wirusach RNA i prawdopodobnie ewoluowały niezależnie od siebie więcej niż jeden raz. Motyw 2A i jemu podobne są obecne jedynie w genomach wirusów (Firth i Brierley, 2012).

3. Regulacja wyboru niekanonicznego kodonu startu

Wykorzystywanie niekanonicznych kodonów w inicjacji translacji jest powszechne u wielu organizmów, między innymi w komórkach ludzkich. Powstałe w ten sposób białka zyskują nowe funkcje, a większość z nich to protoonkogeny, czynniki transkrypcyjne, kinazy czy czynniki wzrostu (Touriol et al., 2003). W zależności od położenia kodonu non-AUG względem kodonu AUG w wyniku niekanonicznej inicjacji translacji może powstawać zarówno skrócone formy białek, gdy niekanoniczny kodon znajduje się poniżej AUG, jak i wydłużone gdy jest powyżej AUG. Jednym z przykładów startu translacji poniżej głównego kodonu inicjującego jest synteza skróconej izoformy ludzkiego białka CAPC (znanego również jako LRRC26). CAPC jest genem wyrażanym w każdej normalnej komórce, natomiast jego skrócona forma (S-CAPC), syntetyzowana z kodonu non-AUG leżącego poniżej kodonu kanonicznego, jest protoonkogenem (Anaganti et al., 2009).

Zdecydowanie więcej jest znanych przypadków, kiedy na skutek startu translacji z kodonów non-AUG powstają izoformy białek wydłużone na końcu N. Profilowanie rybosomów w mysich embrionalnych komórkach macierzystych wykazało, że około 60% przypadków startu translacji zachodzi z kodonów niekanonicznych, w tym ponad 75% stanowią uORF, znajdujące się w rejonach liderowych transkryptów. Z punktu czysto statystycznego wykorzystywanie kodonów niekanonicznych w inicjacji translacji jest częstsze niż kanonicznych, ale nie przekłada się to na faktyczny udział białek syntetyzowanych z non-AUG w komórce, gdyż inicjacja z kodonów innych niż AUG jest znacznie mniej wydajna. Warto podkreślić, że profilowanie rybosomów może dawać pewien procent fałszywie pozytywnych wyników (Ingolia et al., 2011; Kears e i Wilusz, 2017).

Niekanoniczna inicjacja translacji jest istotna w rozwoju chorób neurodegeneracyjnych, takich jak choroba Huntingtona, ataksja rdzeniowo-mózdkowa typu 8, dystrofia miotoniczna typu 1 czy otępienie czołowo-skroniowe. Odpowiada za to szczególnie rodzaj niekanonicznej translacji zwany RAN (ang. *repeat-associated non-AUG translation*), który prowadzi do translacji powtórzeń CAG, CGG i GGGGCC. Mechanizm ten skutkuje powstaniem złogów, które odpowiadają za rozwój choroby. Obecnie wiadomo, że ponad 40 ludzkich chorób neurodegeneracyjnych jest związanych z RAN (Wojciechowska et al., 2014; Cleary et al., 2018), a powtórzenia CAG, CGG i GGGGCC są często celem dla nowych terapeutyków (Disney et al., 2012).

Za kontrolę wyboru kodonu startu translacji odpowiadają czynniki działające *in-trans*, a w szczególności eukariotyczne czynniki inicjujące eIF. Pierwsze badania u drożdży

S. cerevisiae wskazały na udział czynników eIF1, eIF2 oraz eIF5 w tym procesie (Castilho-Valavicius et al., 1990; Huang et al., 1997), podczas gdy kolejne wykazały, że wybór kodonu startu jest silnie regulowany i biorą w nim udział wszystkie eIF obecne w kompleksie preinicjującym (Kearse i Wilusz, 2017).

Jak wielokrotnie wspomniano, inicjacja translacji z kodonów niekanonicznych zachodzi z reguły na niskim poziomie. Największą wydajność wykazują kodony NUG (gdzie N to dowolna zasada), szczególnie CUG i GUG (oba NUG), oraz ACG i AUU (Kolitz et al., 2009; Kearse i Wilusz, 2017; Hinnebusch, et al., 2016; Na et al., 2018). Najrzadziej wykorzystywane są kodony AAG i AGG (Peabody, 1989; Touriol et al., 2003; Kolitz et al., 2009; Kearse i Wilusz, 2017), które w komórkach drożdżowych najprawdopodobniej w ogóle nie inicjują translacji, ponieważ parujący z nimi antykodon zbudowany z samych puryn nie tworzy helisy A i nie jest rozpoznawany przez rybosom jako potencjalne miejsce startu (Chang et al., 2010).

Przeważnie do inicjacji translacji z kodonów niekanonicznych, podobnie jak dla AUG, wykorzystywany jest tRNA^{Met}, a jako pierwszy aminokwas wbudowywana jest metionina (Peabody, 1989; Na et al., 2018), ponieważ eIF2 odpowiadający za przyłączenie inicjującego tRNA ma silne powinowactwo do tRNA^{Met} (Kearse i Wilusz, 2017). Jednak są przypadki użycia innego tRNA, np. w peptydach wiążących MHC klasy I inicjacja translacji z kodonu CUG prowadzi do inkorporacji leucyny jako pierwszego aminokwasu, za co odpowiada czynnik translacyjny eIF2A (Starck et al., 2012).

Stopień wykorzystania niekanonicznych miejsc inicjacji translacji zależy od warunków w jakich znalazła się komórka. Jednym z mechanizmów obronnych komórek w odpowiedzi na stres jest inaktywacja translacji zależnej od kapu, która jest regulowana przez ścieżkę sygnałową mTOR (ang. *mammalian target of rapamycin*) oraz poziom fosforylacji czynnika eIF2α (Holcik i Sonenberg, 2005; Spriggs et al., 2010; Walsh et al., 2013). Dochodzi do tego m.in. podczas infekcji wirusowej w reakcji na dwuniciowy RNA w celu zahamowania replikacji wirusa (Walsh et al., 2013), lub w odpowiedzi na niezwinięte białka (ang. *unfolded protein response*), aby zredukować translację i umożliwić chaperonom poprawne zwinienie wyprodukowanych już białek (Pavitt i Ron, 2012). Także podczas braku substancji odżywczych i tlenu komórki ograniczają translację zależną od kapu w celu zaoszczędzenia energii i aminokwasów (Holcik i Sonenberg, 2005; Spriggs, et al., 2010; Sriram et al., 2018). Redukcja inicjacji translacji zależnej od kapu prowadzi do wykorzystywania alternatywnych szlaków inicjacji translacji, w tym również

do zwiększenia częstości użycia niekanonicznych kodonów startu. Wiele z takich zdarzeń zachodzi na przykład podczas mejozy, kiedy następuje całkowite przebudowanie proteomu (Brar, et al., 2012; Cheng et al., 2018). Badania profilowania rybosomów wskazują na to, że w skali całego genomu tysiące kodonów non-AUG może być wykorzystywane w inicjacji translacji podczas stresu (Ingolia et al., 2009; Lee et al., 2012). Selekcja odpowiednich kodonów inicjujących jest procesem ściśle regulowanym przez czynniki translacyjne takie jak eIF1, eIF1A, eIF2, eIF3 oraz eIF5 (Kearse i Wilusz, 2017). Różnice w ich poziomie, stechiometrii czy aktywności mogą mieć znaczący wpływ na wybór kodonów start. Na przykład zwiększenie poziomu eIF1 prowadzi do silniejszej selekcji kodonów AUG, przez co inicjacja translacji z kodonów non-AUG jest zredukowana (Hinnebusch et al., 2016). Komórki wydają się regulować wybór kodonów startu w odpowiedzi na stres. Dla przykładu, w fibroblastach i komórkach krwi w sytuacji wysokiej gęstości komórek w hodowli, białko c-Myc jest syntetyzowane z położonego powyżej kodonu CUG. Ten efekt jest wynikiem niedoboru metioniny, co skutkuje translacją dłuższej izoformy c-Myc (Sriram et al., 2018).

Różne formy białek syntetyzowane w wyniku inicjacji translacji z kodonów innych niż AUG, mogą zyskiwać dodatkową funkcję, np. odmienną lokalizację niż dla formy kanonicznej. W proteomie ludzkim przykładem może być czynnik wzrostu FGF2, którego podstawowa izoforma znajduje się w cytoplazmie i ulega sekrecji, a wydłużone izoformy, syntetyzowane z kodonów CUG, mają lokalizację jądrową (Arnaud et al., 1999). Podobne zjawisko obserwuje się też dla innych czynników wzrostowych: FGF-3, hck i VEGF (Acland et al., 1990; Lock et al., 1991; Huez et al., 2001; opisane w Touriol et al., 2003). U drożdży *S. cerevisiae* również obserwuje się zmianę lokalizacji formy wydłużonej, powstałej w wyniku inicjacji z kodonu non-AUG, w stosunku do formy podstawowej. Do tej pory opisano cztery białka Grs1, Ala1, Hfa1 i Hyr1, których niekanoniczne izoformy lokalizują się w mitochondriach, podczas gdy kanoniczne są obecne w jądrze lub cytoplazmie (Chang i Wang, 2004; Tang et al., 2004; Soumi et al., 2014; Kritsiligkou et al., 2017). Podobny mechanizm przewidziano bioinformatycznie także dla Bpl1, Faa2 i Sam2 (Soumi et al., 2014). Również białka Hts1 i Vas1 posiadają dwie izoformy białek, wydłużoną mitochondrialną i kanoniczną cytoplazmatyczną, obie powstałe w wyniku translacji z dwóch osobnych kodonów AUG (Natsoulis et al., 1986; Chiu et al., 1992; Chatton et al., 1988). Niewykluczone, że wykorzystanie alternatywnych miejsc startu translacji (ang. *alternative translation initiation start*; aTIS) do syntezy niekanonicznych izoform białek ma wpływ na kształt i różnorodność proteomu mitochondrialnego.

4. Mitochondrium i jego funkcje u drożdży *Saccharomyces cerevisiae*

Mitochondria to pół-autonomiczne organella pochodzenia endosymbiotycznego (Gray et al., 1999). Posiadają własny genom oraz maszynę molekularną umożliwiającą jego replikację, transkrypcję i translację (Gonczarowska-Jorge et al., 2017; D'Souza i Minczuk, 2018). Głównym zadaniem mitochondriów w komórce jest produkcja energii w procesie oddychania komórkowego. U organizmów eukariotycznych, w tym u drożdży *S. cerevisiae*, ostatnim akceptorem elektronów jest tlen. Uwalniana energia pozwala na syntezę ATP, który wykorzystywany jest przez komórkę do różnych reakcji metabolicznych (Klingenberg, 1975). Przez wiele lat sądzono, że mitochondria są niezbędnymi organellami do przeżycia komórki jądrowej, jednak kilka lat temu odkryto w jelitach szynszyli eukariotyczny organizm *Monocercomonoides* sp. PA203 pozbawiony mitochondriów, które zostały wtórnie utracone (Karnkowska et al., 2016). Z kolei drożdże *S. cerevisiae* mogą przeżyć bez funkcjonalnych mitochondriów (Altmann et al., 2007), chociaż ich komórki, zwane *petite*, są znacznie mniejsze od dzikich (Slonimski i Ephrussi, 1949). Drożdże *petite* pozyskują energię tylko poprzez fermentację, dlatego nie są w stanie przeżyć na niefermentowalnym źródle węgla (np. glicerolu) (Chen i Clark-Walker, 2000).

Drożdżowy genom mitochondrialny koduje tylko mitochondrialny rRNA i tRNA oraz kilka białek (Foury et al., 1998; Gonczarowska-Jorge et al., 2017), co stanowi zaledwie ułamek całego proteomu mitochondrialnego. Przeważająca większość białek mitochondrialnych (99%) kodowana jest jądrowo, syntetyzowana w cytoplazmie a następnie transportowana do mitochondriów. Mitochondria otoczone są dwiema błonami: wewnętrzną (ang. *inner mitochondrial membrane*, IMM) i zewnętrzną (ang. *outer mitochondrial membrane*, OMM), między którymi znajduje się przestrzeń międzybłonowa (ang. *intermembrane space*, IMS), a wewnątrz mitochondriów wypełnia macierz mitochondrialna (ang. *mitochondrial matrix*). Transport do mitochondriów odbywa się poprzez szereg transporterów błonowych, które rozpoznają sygnał lokalizacji mitochondrialnej i kierują białka do odpowiednich przedziałów mitochondrialnych (Sokol et al., 2014) (Fig 10).

Sygnałem lokalizacji komórkowej jest znajdująca się na końcu N białka sekwencja MTS (ang. *mitochondrial targeting signal*) kierująca do mitochondriów (von Heijne et al., 1989). MTS zwykle ma długość od 15 do 50 aminokwasów (Wiedemann i Pfanner, 2017), a u drożdży *S. cerevisiae* waha się 6 do 122 aminokwasów (Vögtle et al., 2009; Teixeira i Glaser, 2013). Prekursory białek są transportowane do mitochondriów w formie

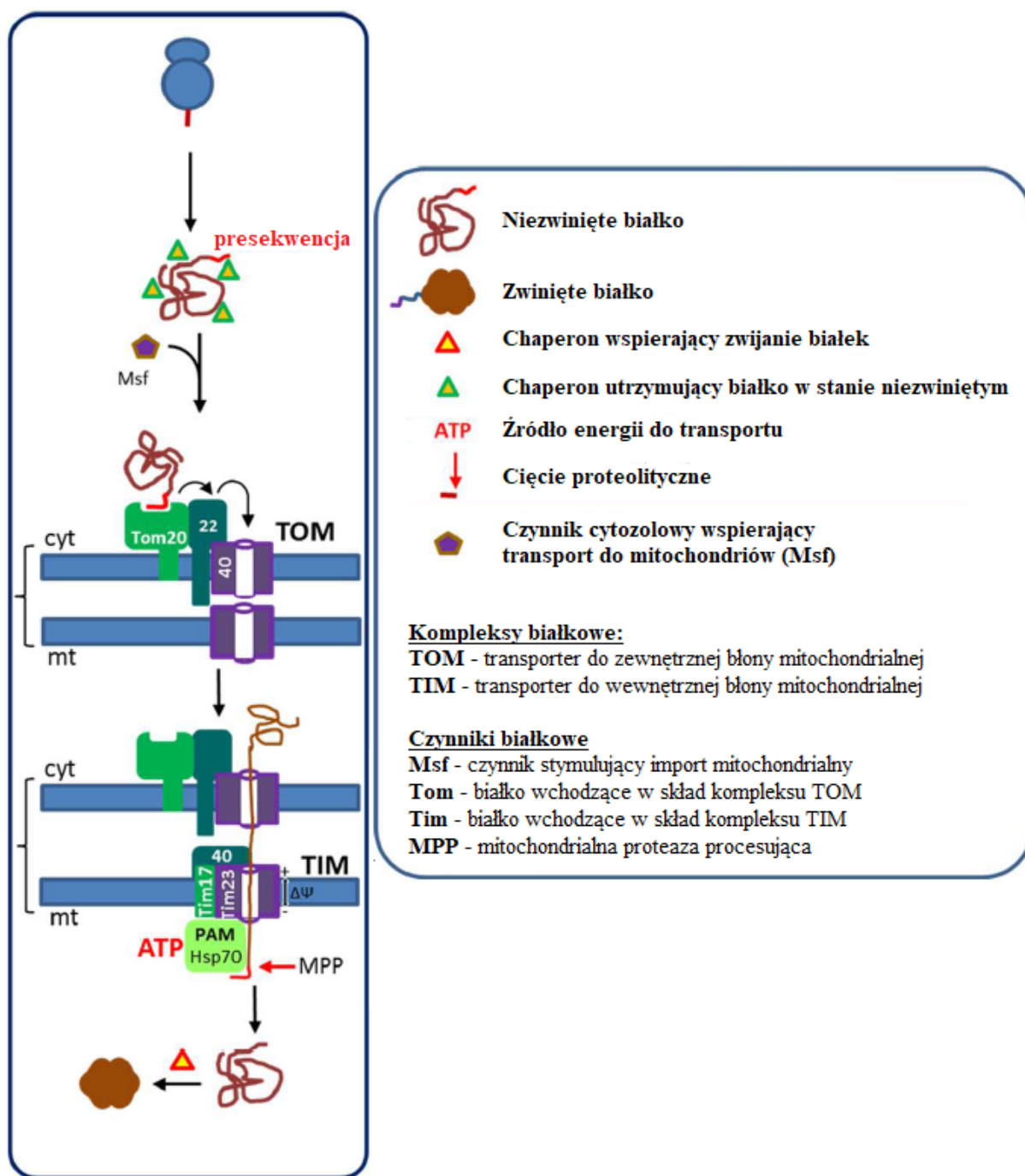


Fig 10. Schemat translokacji białek do mitochondriów. Po zakończeniu translacji białko utrzymywane jest w formie niezwiniełej. Następnie presekwencja na końcu N (MTS) zostaje rozpoznana przez Msf i po przyłączeniu do Tom20 białko transportowane jest przez zewnętrzną, a następnie wewnętrzną błonę mitochondrialną. W trakcie translokacji presekwencja jest odcinana przez MPP, a białko ulega finalnemu zwinięciu w macierzy mitochondrialnej. Figura na podstawie (Kunze i Berger, 2015).

niesfałdowanej, gdzie sekwencja liderowa MTS jest odcinana w macierzy mitochondrialnej przez peptydazy procesujące MPP (ang. *mitochondrial procesing peptidase*), Mas1 i Mas2 u drożdży (Yang et al., 1988), po czym samo białko ulega zwinięciu (Gakh et al., 2002;

Sokol, et al., 2014). Cięcie proteolityczne w mitochondriach wpływa również na stabilność dojrzałych białek w mitochondriach (Vögtle et al., 2009). Istnieje kilka motywów rozpoznawanych przez MPP (Teixeira i Glaser, 2013), chociaż nie zawsze MTS ulega odcięciu. W podobny sposób białka transportowane są do peroksysomów i ER oraz chloroplastów u roślin (Kunze i Berger, 2015).

Sekwencja MTS nie posiada konserwowanej sekwencji aminokwasowej, ale cechuje się określoną budową (von Heijne et al., 1989; Kunze i Berger, 2015), dzięki której można przewidzieć prawdopodobieństwo lokalizacji mitochondrialnej białka za pomocą programów bioinformatycznych takich jak PSORT (Nakai i Horton, 1999), lub bardziej dedykowanych dla MTS: MitoProt II (Claros et al., 1995), TargetP (Emanuelsson et al., 2007) czy MitoFates (Fukusawa et al., 2015). Programy te, analizując najpopularniejsze motywy, są w stanie również przewidzieć miejsce cięcia dla proteaz procesujących. Interesujące jest to, że niektóre białka, pomimo braku przewidzianego MTS, lokalizują się w mitochondriach. Tylko 60% białek mitochondrialnych syntetyzowanych w cytozolu posiada odcinaną presekwencję (Wiedemann i Pfanner, 2017). Znane są wewnętrzne sygnały lokalizacji mitochondrialnej iMTS-Ls (ang. *internal MTS-like signals*), które pełnią takie same funkcje jak MTS, ale są dość słabo określone (Backes et al., 2018).

Według bazy danych genomu *S. cerevisiae*, Saccharomyces Genome Database (SGD), mitochondria drożdży zawierają 1187 białek (Cherry et al., 2012; Gonczarowska-Jorge et al., 2017). Ostatnie prace nad proteomem mitochondrialnym pokazują, że jest on zdecydowanie bardziej zróżnicowany, niż początkowo sądzono. W najnowszym mitoproteomie drożdżowym opracowanym za pomocą spektrometrii mas znajduje się ponad 3400 białek, z których 901 zaklasyfikowano do najwyższej, pierwszej klasy czystości (Morgenstern et al., 2017). Wciąż jednak nie jest wiadome, co jest powodem tak ogromnej różnorodności proteomu mitochondrialnego.

Cel pracy

Celem pracy było zbadanie, czy wykorzystanie niekanonicznych kodonów w inicjacji translacji jest powszechnym zjawiskiem u drożdży *Saccharomyces cerevisiae*.

- Czy powstałe w ten sposób izoformy białek zyskują dodatkowe funkcje, a w szczególności odmienną lokalizację?
- Czy ten proces znacząco wpływa na różnorodność proteomu mitochondrialnego?

Materialy i metody

1. Część bioinformatyczna

1.1. Wykrywanie wydłużonych form białek

W celu wyznaczenia potencjalnych wydłużonych form białek NTE (ang. *N-terminally extension*) wykonano analizę bioinformatyczną, opisaną w Monteuiis, Miścicka et al., 2019.

Geny wysokiego prawdopodobieństwa (ang. *high confidence genes*) wydzielono poprzez odfiltrowanie z puli wszystkich genów tych ORF, które zostały opisane w Saccharomyces Genome Database (SGD) jako „połączone” („merged”), „usunięte” („deleted”) oraz „wątpliwe” („dubious”) które są pozbawione obecności rybosomów w analizie Ribo-seq.

1.2. Analiza GO Term (Gene Ontology)

Analizę GO Term przeprowadzono, używając narzędzi, dostępnych na stronie Saccharomyces Genome Database, dostosowanych do pracy z drożdżami. Do określenia funkcji białek wykorzystano programy GO Term Finder, a do wskazania lokalizacji białek programu GO Slim Mapper.

1.3. Określenie lokalizacji komórkowej

Prawdopodobieństwo lokalizacji mitochondrialnej oraz potencjalne miejsce cięcia przez mitochondrialne proteazy procesujące oszacowano używając programów MitoProt II (Claros et al., 1995), TargetP (Emanuelsson et al., 2007) oraz MitoFates (Fukusawa et al., 2015). Parametry MTS^+ (ang. *MTS gain*) i MTS^- (ang. *MTS loss*) zgodnie ze wzorem:

$$\frac{(N^{MP} - M^{MP}) + (N^{TP} - M^{TP}) + (N^{MF} - M^{MF})}{3}$$

$N^{MP}/N^{TP}/N^{MF}$ oraz $M^{MP}/M^{TP}/M^{MF}$ oznaczają odpowiednio prawdopodobieństwo lokalizacji mitochondrialnej dla formy wydłużonej (N) oraz kanonicznej (M), wyliczonej przez MitoProt II (MP), TargetP (TP) i MitoFates (MF). Jeżeli wynik równania był większy niż 0,3, przypisywano parametr MTS^+ , zaś jeżeli wynik był mniejszy niż -0,3 – MTS^- .

Sygnał lokalizacji jądrowej NLS (ang. *nuclear localisation signal*) dla białka Qri5 wykryto używając programów cNLS Mapper (Kosugi et al., 2009) i NucPred (Brameier et al., 2007).

1.4. Analiza konserwacji

Ludzkie homologi drożdżowych białek wraz z sekwencjami zostały pobrane z bazy ENSEMBL. Alternatywne kodony startu translacji wykryto programem PreTIS (Reuter et al., 2016). Wybrano te aTIS, których użycie tworzyłoby wydłużone na końcu N izoformy białek (analogicznie do ich drożdżowych odpowiedników). MTS oraz parametry MTS^+ i MTS^- zostały zdefiniowane jak poprzednio. Dana izoforma została uznana za mitochondrialną, jeżeli przewidziany dla niej MTS przez MitoProt II, TargetP lub MitoFates wynosił powyżej 0,3.

2. Część molekularna

2.1. Podstawowe techniki biologii molekularnej

Podstawowe techniki biologii molekularnej takie jak elektroforeza w żelu agarozowym, izolacja plazmidowego i genomowego DNA, PCR, trawienie enzymami restrykcyjnymi, ligacja itd., były przeprowadzane zgodnie z zaleceniami producenta odczynników oraz według Sambrook et al., 1989.

2.2. Hodowla drożdży

Drożdże hodowano na podłożu pełnym: 1% ekstrakt drożdżowy (Difco Laboratories), 2% bacto-pepton (Difco Laboratories) z odpowiednim źródłem węgla: 2% glukozy (YPD), 3% glicerolu (YPG) lub 2% galaktozy (YPGal). Pożywka YPG była dodatkowo doprowadzana do pH=5,4 za pomocą stężonego kwasu solnego. Do hodowli na podłożu minimalnym stosowano podłoże syntetyczne o składzie 0,67% drożdżowa baza azotowa bez aminokwasów (Difco Laboratories), siarczan amonu, mieszanina aminokwasów bez uracylu, leucyny lub histydy (w zależności od układu eksperymentalnego) w stężeniu zalecanym przez producenta (Qbiogene), glukoza 2%. Hodowle płynne prowadzono w łaźni wodnej z wytrząsaniem. Podłoże stałe zawierało dodatkowo 2% agar. Wszystkie hodowle inkubowano w 30°C. Pełna lista szczepów, użytych w pracy, została zawarta w Tabeli 1.

Komórki drożdżowe do dłuższego przechowywania mrożone były w próbkach do głębokiego mrożenia w ciekłym azocie w pożywce YPD z 30% glicerolem i przechowywane w temperaturze -80°C.

Szczep	Genotyp	Źródło
BMA38	<i>MATa ade2-1 his3Δ200 leu2-3,112 trp1Δ1 ura3-1</i>	(Mayes et al., 1999)
BY4741	<i>MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0</i>	Euroscarf
SC0000	<i>MATa ade2 arg4 leu2-3,112 trp1-289 ura3-52</i>	Euroscarf
<i>trz1Δ/TRZ1</i>	<i>MATa/MATa; ura3Δ0/ura3Δ0; leu2Δ0/leu2Δ0; his3Δ1/his3Δ1; met15Δ0/MET15; LYS2/lys2Δ0; TRZ1/TRZ1::kanMX4</i>	Euroscarf
<i>TRZ1-TAP</i>	jak SC0000 ale <i>YKR079c::TAP-KIURA3</i>	Euroscarf
<i>MTR4-TAP</i>	jak SC0000 ale <i>YJL050w::TAP-KIURA3</i>	Euroscarf
<i>GPP1-TAP</i>	jak SC0000 ale <i>YIL053w::TAP-KIURA3</i>	Euroscarf
<i>GPP2-TAP</i>	jak SC0000 ale <i>YER062c::TAP-KIURA3</i>	Euroscarf
<i>TRX2-TAP</i>	jak SC0000 ale <i>YGR209c::TAP-KIURA3</i>	Euroscarf
<i>PMT2-TAP</i>	jak SC0000 ale <i>YAL023c::TAP-KIURA3</i>	Euroscarf
<i>LSM1-TAP</i>	<i>MATa ade2 his3 leu2 trp2 ura3 YJL124c::TAP-TRP1</i>	(Kufel et al., 2004)
<i>ATG1-TAP</i>	jak SC0000 ale <i>YGL180w::TAP-KIURA3</i>	Euroscarf
<i>IOC4-TAP</i>	jak SC0000 ale <i>YMR044w::TAP-KIURA3</i>	Euroscarf
<i>NOP9-TAP</i>	jak SC0000 ale <i>YMR044w::TAP-KIURA3</i>	Euroscarf
<i>ADH4-TAP</i>	jak SC0000 ale <i>YGL256w::TAP-KIURA3</i>	Euroscarf

<i>LSM4-TAP</i>	jak SC0000 ale <i>YER112w::TAP-KIURA3</i>	Euroscarf
<i>TRM3-TAP</i>	jak SC0000 ale <i>YDL112w::TAP-KIURA3</i>	Euroscarf
<i>HYR1-TAP</i>	jak SC0000 ale <i>YIR037w::TAP-KIURA3</i>	Euroscarf
<i>FUR1-TAP</i>	jak SC0000 ale <i>YHR128w::TAP-KIURA3</i>	Euroscarf
<i>MUM2-TAP</i>	jak SC0000 ale <i>YBR057c::TAP-KIURA3</i>	Euroscarf
<i>LSM1-MORF</i>	<i>MATa pep4-3, his4-580, ura3-52, leu2-3, 112 [2μ URA3 GAL-LSM1-His₆-HA-(3C)-ZZ]</i>	(Gelperin et al., 2005); OpenBiosystems
<i>TRZ1-MORF</i>	<i>MATa pep4-3, his4-580, ura3-52, leu2-3, 112 [2μ URA3 GAL-TRZ1-His₆-HA-(3C)-ZZ]</i>	(Gelperin et al., 2005); OpenBiosystems
<i>GAL1::3HA-MTR4</i>	<i>MATa ade2 his3 leu2 trp2 ura3 GAL1::3HA-MTR4::HIS5sp</i>	(Torchet et al., 2002)
<i>GAL-SmE</i>	<i>MATa ura3-52 lys2-801 ade2-110 trp1-D63 his3D200 SME1::HIS3, Pgal1:SmE (CEN TRP1)</i>	(Camasses et al., 1998)
<i>lsm1Δ</i>	jak BMA38 ale <i>LSM1::TRP1</i>	(Mayes et al., 1999)
<i>lsm6Δ</i>	jak BMA38 ale <i>LSM6::HIS3</i>	(Mayes et al., 1999)
<i>atg1Δ</i>	jak BY4741 ale <i>ATG1::kanMX4</i>	Euroscarf
<i>ioc4Δ</i>	jak BY4741 ale <i>IOC4::kanMX4</i>	Euroscarf
<i>mum2Δ</i>	jak BY4741 ale <i>MUM2::kanMX4</i>	Euroscarf
<i>trm3Δ</i>	jak BY4741 ale <i>TRM3::kanMX4</i>	Euroscarf
<i>trx2Δ</i>	jak BY4741 ale <i>TRX2::kanMX4</i>	Euroscarf
<i>adh4Δ</i>	jak BY4741 ale <i>ADH4::kanMX4</i>	Euroscarf
<i>gpp1Δ</i>	jak BY4741 ale <i>GPP1::kanMX4</i>	Euroscarf
<i>gpp2Δ</i>	jak BY4741 ale <i>GPP2::kanMX4</i>	Euroscarf
<i>hyr1Δ</i>	jak BY4741 ale <i>HYR1::kanMX4</i>	Euroscarf

Tabela 1. Lista szczepów drożdżowych użytych w pracy. Większość szczepów, zawierających białka znakowane na końcu C znacznikiem TAP-tag, zakupiono z kolekcji szczepów firmy Euroscarf (Frankfurt nad Menem, Niemcy).

2.3. Konstrukcja plazmidów

Wszystkie plazmidy użyte w pracy zostały wymienione w Tabeli 2.

Plazmid	Źródło
p416-ADH	(Mumberg et al., 1995)
p415-ADH	(Mumberg et al., 1995)
p415-ADH-TRZ1 ^{ATG}	Ta praca
p416-ADH-TRZ1 ^{TTC}	Ta praca
p416-ADH-TRZ1 ^{TTC} -TAP	Ta praca
p415-ADH-TRZ1 ^{TTC/GCG}	Ta praca
p416-ADH-TRZ1 ^{TTC/GCG} -TAP	Ta praca

Tabela 2. Lista plazmidów użytych w pracy. Metoda konstrukcji plazmidów, utworzonych na potrzeby tej pracy, opisana poniżej.

Plazmidy, wyrażające różne warianty białka Trz1 pod konstytutywnym promotorem genu *ADH1*, otrzymano poprzez klonowanie odpowiednich sekwencji genu *TRZ1* do wektorów ekspresyjnych (Mumberg et al., 1995) za pomocą enzymu SmaI oraz metody jednoetapowego klonowania niezależnego od ligacji (ang. *one-step sequence- and ligation-*

independent cloning, SLIC) (Jeong et al., 2012). Poprawność przygotowania plazmidów sprawdzano poprzez analizę map restrykcyjnych.

Wektory do których miała zostać wklonowana wstawka, zostały strawione enzymem FastDigest SmaI (Thermo Fisher) przez 1 godzinę w 37°C w buforze dołączonym przez producenta. Po tym czasie enzym zdezaktywowano w 80°C przez 5 min i bezpośrednio użyto do reakcji SLIC (bez oczyszczania). Nadmiar trawionego plazmidu był mrożony w -20°C i używany w kolejnych reakcjach. Przygotowano następującą mieszaninę reakcyjną: 100 ng wektora, wstawka w stosunku do wektora molowo 1:4, 1 µl buforu do trawień Green lub FastDigest (Thermo Fisher), 1 µl BSA (0,1 mg/ml; NEB) oraz woda dejonizowana do objętości 10 µl. Po wymieszaniu dodano 0,5 µl Polimerazy T4 DNA (NEB) i przeprowadzano reakcję przez 2,5 min w temperaturze pokojowej. Reakcję zatrzymywano poprzez 10-minutową inkubację w lodzie. Następnie całość mieszaniny reakcyjnej użyto do transformacji bakterii *Escherichia coli*, szczep MH1 (Tabela 4).

Sekwencje wariantów genu *TRZ1* zostały namnożone przy użyciu techniki PCR na matrycy genomowego DNA ze szczepu dzikiego BY4741 lub szczepu *TRZ1-TAP* w zależności od układu eksperymentalnego, za pomocą odpowiednich starterów, opisanych w Tabeli 3. Po reakcji produkty PCR były oczyszczane za pomocą zestawu Wizard® SV Gel and PCR Clean-Up System (Promega) zgodnie z przepisem, dołączonym przez producenta. Plazmidy namnażane były w bakteriach *Escherichia coli*, szczep DH5α (Tabela 4).

Mutacje punktowe zmieniające kanoniczny kodon startu translacji ATG na GCG w wariacie *TRZ1^{TTC/GCG}-TAP*, zostały wprowadzone przy użyciu techniki PCR poprzez namnożenie fragmentów genu *TRZ1* stosując startery F-TRZ1-TTC i R-TRZ1-GCG oraz F-TRZ1-GCG i R-TRZ1-NdeI (Tabela 3) na matrycy genomowego DNA ze szczepu BY4741. Otrzymane produkty reakcji zostały połączone za pomocą techniki PCR we wstawkę F-TRZ1-TTC:TRZ1-GCG:R-TRZ1-NdeI, którą strawiono enzymem restrykcyjnym FastDigest BamHI (Thermo Fisher), zgodnie z zaleceniami producenta. Wektor p416-*ADH-TRZ1^{TTC}-TAP* przygotowano do ligacji poprzez jednoczesne trawienie enzymem FastDigest BamHI oraz defosforylację alkaliczną fosfatazą CIAP (Thermo Fisher) zgodnie z zaleceniami producenta, po czym wligowano strawioną wstawkę F-TRZ1-TTC:TRZ1-GCG:R-TRZ1-NdeI za pomocą ligazy T4 DNA (Thermo Fisher) według standardowego protokołu. Mutacje punktowe zostały potwierdzone za pomocą sekwencjonowania.

Nazwa	Opis	Sekwencja
F-Trz1-ATG	Namnożenie genu <i>TRZ1</i> począwszy od kanonicznego kodonu ATG	TCTAGAACTAGTGGATCCCCCATGTTC ACATTTATACCCATCACCC
F-Trz1-TTC	Namnożenie genu <i>TRZ1</i> począwszy od niekanonicznego kodonu TTC	TCTAGAACTAGTGGATCCCCCTTCAAA CTGTCTGAATATACTACTAAATAG
R-Trz1	Namnożenie genu <i>TRZ1</i> bez znacznika TAP-tag	ATATCGAATTCCTGCAGCCCATTTTTC TTGTGTTTCTTAAGTTTGACTTC
R-Trz1-TAP	Namnożenie genu <i>TRZ1</i> ze znacznikiem TAP-tag	ATATCGAATTCCTGCAGCCCTCAGGTT GACTTCCCCGCG
F-Trz1-GCG	Starter do wprowadzenia mutacji	TGGAAGGTTTTCAACGGCGTTCACATT TATA
R-Trz1-GCG	Starter do wprowadzenia mutacji	ATGGGTATAAATGTGAACGCCGTTG
F-Trz1-mutacja-ATG	Starter do potwierdzenia obecności kodonu ATG w genie <i>TRZ1</i>	AGTCCTGGAAGGTTTTCAACGAT
F-Trz1-mutacja-GCG	Starter do potwierdzenia obecności mutacji GCG w genie <i>TRZ1</i>	TCCTGGAAGGTTTTCAACGGC
F-Trz1-delta	Sprawdzenie delecji genu <i>TRZ1</i> w szczepie <i>trz1</i> Δ	GGCATCGCATAATCTATAACTGAC
R-Trz1-delta	Sprawdzenie delecji genu <i>TRZ1</i> w szczepie <i>trz1</i> Δ	CCTTTTTGAAAACATATCGAAGACGC
R-Trz1-NdeI	Starter pomocniczy do przygotowania wariantu Trz1 ^{TTC/GCG}	ATTGTGGGATATCATATGATTACTTT ACCAT
F-Trz1-Sekw	Starter do sekwencjonowania w celu sprawdzenia wprowadzonej mutacji	CATACAGAAAGGGACC

Tabela 3. Lista starterów użytych w pracy. Dokładny opis w tekście.

Szczep	Genotyp	Zastosowanie
E. coli MH1	araD139, ΔlacX74, galU, galK, hsdR, rpsL	Biorca reakcji SLIC
E. coli DH5α	F– Φ80lacZΔM15, Δ(lacZYA-argF), U169, recA1, endA1, hsdR17 (rK–, mK+), phoA, supE44 λ– thi-1, gyrA96, relA1	Namnażanie gotowych konstruktów

Tabela 4. Lista szczepów bakteryjnych użytych w pracy. Stosowano bakterie chemokompetentne, przygotowane i transformowane jak w Chang et al., 2017.

Do potwierdzenia lokalizacji wariantu Trz1^{ATG}-TAP użyto szczepu Trz1-MORF (Gelperin et al., 2005), który zawiera konstrukt wyrażający białko Trz1-TAP syntetyzowane z kanonicznego kodonu startu AUG pod kontrolą indukowanego promotora *GAL1*.

2.4. Transformacja drożdży

Do transformacji drożdży przygotowanymi konstruktami plazmidowymi wykorzystano metodę elektroporacji. Pobrano 1,5 ml hodowli nocnej, zwirowano (5 min, 3000xg) i przepłukano dwukrotnie buforem do transformacji (20 mM HEPES, 1 M sorbitol). Następnie zawieszono pelet w 50 µl w tym samym buforze, dodano 1 µg plazmidu i inkubowano przez 10 minut na lodzie. Tak przygotowaną mieszaninę przeniesiono

do schłodzonych kuwet elektroporacyjnych i poddano impulsowi elektrycznemu o napięciu 1600 V za pomocą elektroporatora Eppendorf Electroporator 2510.

2.5. Sporulacja i rozkładanie tetrad

W celu przygotowania szczepów wyrażających różne warianty Trz1, przeprowadzono transformację szczepu diploidalnego *TRZ1/trz1Δ* odpowiednimi konstruktami plazmidowymi. Następnie transformanty hodowano na pożywce presporulacyjnej GNA (3% bulion odżywczy (Difco), 1% ekstrakt drożdżowy, 5% glukoza, 2% agar) przez dwa dni w temperaturze 30°C. Po tym czasie przeniesiono je na pożywkę sporulacyjną (1% octan potasu, 0,1% ekstrakt drożdżowy, 0,05% glukoza, 2% agar oraz niezbędne aminokwasy, zależnie od selekcji pokarmowej wyrażonej na plazmidach) i sporulowano przez 5-7 dni w temperaturze 30°C. Po tym czasie tetrady były rozkładane przy użyciu mikromanipulatora SINGER Instruments MSM System, dostępnego w Instytucie Biochemii i Biofizyki PAN. Wyselekcjonowane tetrady sprawdzano pod kątem oporności na G418 (obecność mutacji *trz1Δ* na chromosomie) oraz selekcji pokarmowej -Ura, -Leu lub -His (obecność plazmidu).

Plazmidy *p416-ADH-TRZ1^{TTC}-TAP* oraz *p416-ADH-TRZ1^{TTC/GCG}-TAP* zostały usunięte z komórek po hodowli na szalkach zawierających kwas 5' fluoroorotowy (5'-FOA).

2.6. Frakcjonowanie komórek drożdżowych

Do frakcjonowania komórek drożdżowych użyto zmodyfikowanego protokołu izolacji mitochondriów z drożdży (Meisinger et al., 2000; Meisinger et al., 2006).

Drożdże hodowano przez noc z wytrząsaniem w 30°C. Następnego dnia, używając przygotowanej uprzednio prekultury, zaszczipiono 100 ml pożywki YPD do początkowego OD₆₀₀ ~0,01 i hodowano przez noc jak poprzednio do OD₆₀₀ ~1-2. Zbierano 60 OD poprzez wirowanie w probówkach typu falcon o objętości 50 ml i płukano dwukrotnie wodą dejonizowaną. Tak otrzymywane pelety mrożono w -80°C (do użycia późniejszym terminie) lub -20°C (do użycia następnego dnia) bądź przystępowano do dalszej procedury.

Świeże bądź mrożone osady drożdżowe zawieszano w 6 ml buforu DTT (10 mM DTT, 100 mM Tris-SO₄ pH=9,4) i inkubowano z wytrząsaniem przez 15 min w 30°C. Następnie komórki były jednokrotnie płukane buforem do trawienia zymoliazą (1,2 M sorbitol, 20 mM Kpi pH=7,4), a potem zawieszone w 6 ml tego samego buforu. Po dodaniu 30 µl Zymoliazы T20 (10 mg/ml), komórki inkubowano z wytrząsaniem przez 30 min w 30°C. Po tym czasie sferoplasty zwirowano (5 min, 3000xg, 25°C). Od tego momentu

używano schłodzonych buforów i wszystkie procedury przeprowadzano na lodzie. Sferoplasty przepłukano jednokrotnie buforem do homogenizacji (0,6 M sorbitol, 10 mM Tris-HCl pH=7,4 z dodatkiem inhibitora proteaz Complete ULTRA Tablets firmy Roche). Po zawieszeniu w 6 ml tego samego buforu, sferoplasty ucierano w homogenizatorze Pottera ze szklanym tłokiem (4-5 pociągnięć). Następnie homogenaty zostały zwirowane (5 min, 3000xg, 4°C) i rozdzielone na cztery frakcje. 1/6 supernatantu zostało przeniesione do probówki typu eppendorf (ekstrakt całkowity, frakcja T), natomiast 1/6 i 2/3 po przeniesieniu do probówek typu eppendorf zostały zwirowane przez 15 min przy prędkości 18000xg w temperaturze 4°C (odpowiednio frakcje mitochondrialna M i zagęszczona mitochondrialna D). Supernatant po wirowaniu frakcji M został przeniesiony do czwartej probówki (frakcja cytozolowa C), natomiast supernatant frakcji D odrzucono. Pelety mitochondrialne zawieszono w 130 µl (frakcja M) i 60 µl (frakcja D) buforu do próbek (6 M mocznik, 6% SDS, 50 mM DTT, 0,01% błękit bromofenolowy, 125 mM Tris-HCl, pH=6,8), następnie zdenaturowano w 65°C z wytrząsaniem przez 20 minut.

2.7. Precypitacja białek przy użyciu kwasu trójchlorooctowego (TCA)

Frakcje ekstraktu całkowitego (T) oraz cytozolowa (C) zostały poddane procedurze precypitacji białek z użyciem kwasu trójchlorooctowego (TCA). Do frakcji T i C dodano 1/10 objętości 100% TCA (do końcowego stężenia 10% w próbce), inkubowano na lodzie przez 20 min, po czym zwirowano (15 min, 18000xg, 4°C). Następnie odrzucono supernatant, a do osadu dodano 800 µl zimnego acetonu, po czym ponownie zwirowano jak wcześniej. Zlano supernatant, pelety wysuszone w temperaturze pokojowej i dodano 130 µl buforu do próbek (skład jak wcześniej), następnie zdenaturowano w 65°C z wytrząsaniem przez 20 minut.

2.8. Deglikozylacja endoglikozydazą H (EndoH)

Do deglikozylacji frakcji cytozolowej i mitochondrialnej po frakcjonowaniu szczepu Mtr4-TAP użyto endoglikozydazy H (EndoH; New England Biolabs) zgodnie z zaleceniami producenta.

Deglikozylacja frakcji cytozolowej (C). Do 154 µl frakcji cytozolowej (C) przed strącaniem dodano 17 µl 10x buforu do denaturacji glikoprotein (GDB, Glycoprotein Denaturing Buffer 10x, dołączony przez producenta) po czym przeprowadzano denaturację przez 10 min w bloku grzewczym w temperaturze 100°C. Następnie dodano 20 µl 10x GlycoBuffer3

(dołączony przez producenta) i 9 μ l EndoH. Trawienie przeprowadzono przez 1 godzinę w temperaturze 37°C, dodano 800 μ l H₂O i strącono białka z użyciem TCA jak wyżej.

Deglikozylacja frakcji mitochondrialnej zagęszczonej (D). Do 20 μ l peletu zawieszonego w buforze SM (250 mM cukroza, 10 mM MOPS-KOH pH=7,2) dodano 3 μ l 10x GDB oraz 7 μ l H₂O. Po denaturacji dodano 4 μ l 10x GlycoBuffer3 i 1,8 μ l EndoH. Denaturację oraz trawienie przeprowadzano równocześnie i w tych samych warunkach co w przypadku frakcji cytozolowej. Reakcję dopełniono do 1 ml wodą miliQ i strącono białka z użyciem TCA.

2.9. Defosforylacja białkową fosfatazą faga Lambda (LPP)

Defosforylację frakcji cytozolowej i mitochondrialnej po frakcjonowaniu szczepu Mtr4-TAP wykonano używając białkowej fosfatazy faga Lambda (LPP, New England Biolabs) zgodnie z zaleceniami producenta.

Defosforylacja frakcji cytozolowej (C). Do 154 μ l frakcji cytozolowej (C) przed strącaniem, dodano 20 μ l buforu 10x PMP (dołączony przez producenta), 20 μ l 10x MnCl₂ i 6 μ l LPP. Reakcję inkubowano przez 15 minut w temperaturze 30°C, dopełniono do 1 ml wodą i strącono białka przy użyciu TCA.

Defosforylacja frakcji mitochondrialnej zagęszczonej (D). Do 20 μ l peletu zawieszonego w buforze SM (250 mM cukroza, 10 mM MOPS-KOH pH=7,2) dodano 4 μ l 10x PMP, 4 μ l 10x MnCl₂ i 1,2 μ l LPP. Następnie postępowano tak samo jak przy frakcji cytozolowej.

2.10. Elektroforeza białek SDS-PAGE

Białka rozdzielano w denaturujących żelach poliakrylamidowych (akrylamid/bisakrylamid w stosunku 37,5:1, Sigma) w obecności SDS w systemie Laemmli'ego. Procentowość żeli akrylamidowych była dobierana w zależności od masy cząsteczkowej badanych białek (8%, 10%, 12% lub 15%). W celu określenia czystości frakcjonowania ekstrakty były rozdzielane w żelach 12%. Na żele nanoszono równe objętości wszystkich frakcji (po 20 μ l). Rozdział elektroforetyczny przeprowadzano przy stałym napięciu 130 V aż do opuszczenia przez barwnik granicy żelu w buforze Tris-glicyna-SDS (25 mM Tris-Base, 0,192 M glicyna, 0,1% SDS). Następnie przeprowadzano elektrotransfer mokry na membranę PVDF (Bio-Rad) w buforze do transferu (25 mM Tris-Base, 0,192 M glicyna, 20% metanol) przy napięciu 100 V przez 1-1,5 godziny lub 30V przez noc w chłodni.

2.11. Analiza western blot

Membrany po transferze blokowano przez noc w 5% roztworze odtłuszczonego mleka w PBST w temperaturze 4°C. W celu wykrycia znacznika TAP-tag stosowano przeciwciała PAP (ang. *peroxidase-anti-peroxidase*, Sigma, P1291) w rozcieńczeniu 1:3000. Inkubację z przeciwciałami przeprowadzano w temperaturze pokojowej przez 3 godziny. Po tym czasie odpłukano bufor blokujący oraz nadmiar przeciwciał poprzez trzykrotną inkubację z PBST (137 mM NaCl, 27 mM KCl, 10 mM Na₂HPO₄, 1,8 mM KH₂PO₄, 0,01% TWEEN 20) przez 5 minut. Do detekcji wykorzystano zestaw do wzmocnionej chemiluminescencji ECL (Clarity Western ECL Substrate, Bio-Rad) zgodnie z zaleceniami producenta. Sygnał utrwalono na kliszach fotograficznych Agfa. W celu wykrycia znacznika HA-tag (ang. *human influenza hemagglutinin*) stosowano specyficzne przeciwciała α -HA (Roche, 12CA5) w stężeniu 1:1000 w 5% odtłuszczonego mleka w PBST przez 1 godzinę w temperaturze pokojowej. Następnie przepłukano membrany trzykrotnie po 5 min roztworem PBST, po czym inkubowano z przeciwciałami drugorzędowymi, α -mysimi IgG w stężeniu 1:25000 (Thermo Fisher, 31430) przez 1 godzinę w 5% odtłuszczonego mleka w PBST. Wywoływano tak jak przy wykrywaniu znacznika TAP-tag. Do immunodetekcji białek Ilv2 i Sgt2 po blokowaniu membranę przecięto wzdłuż linii na wysokości 55 kDa. Obie części inkubowano w tym samym czasie z króliczymi poliklonalnymi przeciwciałami α -Ilv2 (fragment powyżej 55 kDa) lub α -Sgt2 (fragment poniżej 55 kDa) w rozcieńczeniach 1:500 w 5% roztworze mleka w TBST (150 mM NaCl, 20 mM Tris-HCl pH=7,4, 0,01% TWEEN 20) przez 1 godzinę w temperaturze pokojowej. Następnie filtry przepłukano trzykrotnie roztworem PBST (5 min), a potem przeprowadzono inkubację z przeciwciałami drugorzędowymi, α -króliczymi IgG (Sigma, A0545) sprzężonymi z peroksydazą chrzanową w rozcieńczeniu 1:10000 w 5% roztworze odtłuszczonego mleka w PBST przez godzinę w temperaturze pokojowej. W takich samych warunkach przeprowadzano detekcję białka Trr1, używając do detekcji króliczych poliklonalnych przeciwciał α -Trr1 w rozcieńczeniu 1:500. Do detekcji białek Ilv2, Sgt2 i Trr1 przygotowano substrat: do 5 ml 100 mM Tris-HCl pH=8,8 dodano 12,5 μ l 90 mM kwasu kumarowego (Sigma), 30 μ l 250 mM luminolu (Sigma) oraz 3 μ l 30% H₂O₂ (Sigma). Tak jak wcześniej, sygnał utrwalono na kliszach fotograficznych Agfa.

2.12. Ekstrakcja małych RNA z drożdży

Frakcję małych RNA z drożdży wyizolowano zmodyfikowaną metodą wg (Wu et al., 2013) z ok. 10 OD drożdży zbieranych przy gęstości optycznej OD₆₀₀~0,4. Pelet przepłukano

sterylną wodą miliQ i zamrożono w ciekłym azocie. Następnie osad zawieszono w 400 μ l zimnego TSE (10 mM EDTA, 500 mM NaCl, 100mM Tris-HCl pH=7,5), dodano 400 μ l kwaśnego fenolu nasyconego TSE (TSE:fenol 1:1) i inkubowano 1 h w bloku grzewczym w temperaturze 37°C, z wytrząsaniem 200-300 rpm. Co 15 min próbki mieszano poprzez obracanie. Następnie próbki przełożono na lód na 30 min, również mieszając co 15 min przez obracanie. Po tym czasie próbki zworteksowano przy niskich obrotach przez 2 minuty i zwirowano 6 min 15 000 rpm w temperaturze pokojowej. Fazę wodną ekstrahowano dwukrotnie fenolem nasyconym TSE, po czym strącono 3 objętościami 98% etanolu w temperaturze -80°C przez noc. Następnego dnia wirowano 20 min 15 000 rpm w wirówce schłodzonej do 4°C, po czym dwukrotnie przepłukano 70% etanolem. Osad wysuszono i zawieszono w 20 μ l H₂O.

2.13. Analiza northern blot

Do rozdzielania RNA używano 8% żeli poliakrylamidowych (akrylamid/bisakrylamid w stosunku 19:1 Sigma) w roztworze 8 M mocznika i w buforze 1x TBE (89 mM Tris base, 89 mM kwas borowy, 2 mM EDTA). Na żel nakładano 10 μ g frakcji małych RNA w buforze do nakładania próbek (98% formamid (Sigma); 0,025% cyjanian ksylenu; 0,025% błękit bromofenolowy) w stosunku 1:1. Przed nałożeniem na żel próbki zostały zdenaturowane w temperaturze 95°C przez 5 minut. Rozdział elektroforetyczny przeprowadzano przy stałej mocy prądu 15W, po czym rozdzielony RNA przeniesiono na membranę nylonową (Hybond N⁺, Amersham Biosciences) za pomocą elektrotransferu mokrego w buforze 0,5x TBE przy stałym natężeniu prądu 10 mA przez noc. Po zakończeniu transferu RNA utrwalano na membranie przez naświetlanie promieniami UV (1200 x 100 μ J/cm²) w komorze sieciującej Crosslinker UVC500, Hoefer.

Prehybrydyzację membran przeprowadzano w piecu hybrydyzacyjnym w temperaturze 37°C przez 30 min w 5 ml buforu hybrydyzacyjnego (5x Bufor Denhardta (0,1% fikor 400; 0,1% poliwinylpirolidon; 0,1% BSA), 6x SSPE (1 M NaCl; 6 mM EDTA; 60 mM bufor fosforanowy pH = 7,4), 0,5% SDS). Następnie w tym samym buforze i warunkach inkubacji prowadzono hybrydyzację przez noc w obecności 10 pmoli specyficznych sond oligonukleotydowych (Tabela 5), wyznakowanych radioaktywnie ³²P- γ ATP (Hartmann Analytics) z użyciem 1 U kinazy polinukleotydowej PNK (Thermo Fisher Scientific) według zaleceń producenta (30 min w temperaturze 37°C). Nadmiar sondy po hybrydyzacji odpłukano poprzez trzykrotną inkubację membrany z buforem 6x SSPE przez 15 min w 37°C

37°C. Ekspozycję przeprowadzano z użyciem ekranu Storage Phosphor Screen BAS-IP MS (Fuji Lifesciences) i skanowano przy użyciu skanera Typhoon FLA 9000 (GE Healthcare).

Starter	Nazwa	Sekwencja
W197	tRNA _{IV} ^{LysI}	GTGTTATTAGCCCTGTAGG
W213	mt-tRNA ^{His}	ACCATTGAAATATATTCACC
W223	mt-tRNA ^{Tyr}	CAACTCCAACACTACCAACATTG
W229	tRNA _{VII} ^{Phe}	GTATGGCTTATTATGCGAATTC

Tabela 5. Lista sond użytych w pracy. Na podstawie Skowronek et al., 2014.

Wyniki

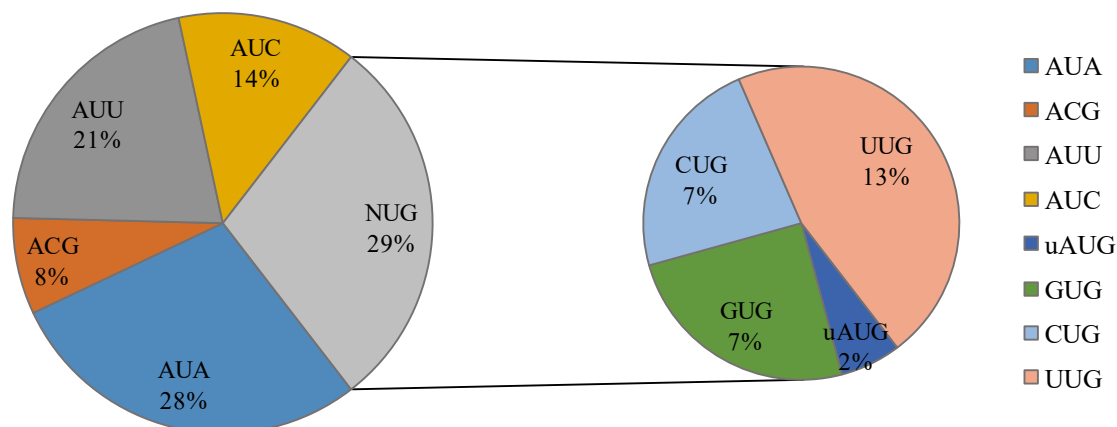
1. Wykrywanie potencjalnych alternatywnych miejsc inicjacji translacji

Do tej pory opisano zaledwie pięć przypadków białek drożdżowych posiadających zarówno izoformy kanoniczne, inicjowane z kodonu AUG, oraz wydłużone, powstałe w wyniku startu translacji z kodonu non-AUG, położonego powyżej opisanego miejsca startu (Chang i Wang, 2004; Tang et al., 2004; Soumi et al., 2014; Kritsiligkou et al., 2017; Heublein et al., 2019). Aby sprawdzić, czy takie przypadki występują powszechnie w genomie drożdżowym, wykonano wyliczenia bioinformatyczne. Jako możliwe alternatywne miejsca startu translacji (aTIS) wybrano kodony spokrewnione z AUG (CUG, UUG, GUG, ACG, AUC, AUA i AUU), które różnią się od niego jednym nukleotydem i mogą służyć do inicjacji translacji u drożdży (Chang et al., 2010; Hinnebusch et al., 2016). Do analiz dołączono też kodon uAUG (*upstream AUG*). Wyliczone przez algorytm wydłużenia na końcu N nazwano NTE (ang. *N-terminal extension*), a zawierające je białka – białkami NTE. Algorytm wyszukiwania regionów NTE został opracowany przez mgr Lounisa Zenad, a udoskonalony i przeprowadzony przez mgr Michała Świrskiego (Monteuuis et al., 2019). Pozostałe analizy z wykorzystaniem powstałej bazy danych przygotowała autorka tej pracy.

Analizie poddano 6 664 drożdżowych ORF opisanych w bazie SGD, a zastosowanie algorytmu ujawniło, że 4 457 z nich może zyskiwać jedno i więcej NTE, dając łącznie 12 953 wydłużonych izoform. Z uzyskanej bazy danych odfiltrowano część genów, tworząc zbiór tzw. genów wysokiego prawdopodobieństwa (ang. *high confidence genes*), na których oparto dalsze analizy (patrz Materiały i Metody). W ten sposób otrzymano 6 110 genów, z których 4 065 zyskuje NTE (łącznie 11 607 izoform) (Tabela 6).

	Cały genom	Geny wysokiego prawdopodobieństwa
Geny	6 664	6 110
Geny NTE	4 457	4 065
Izoformy NTE	12 953	11 607
Geny bez NTE	2 207	2 045
Białka MTS⁺	565	517
Izoformy MTS⁺	1 286	1 199
Białka MTS⁻	233	206
Izoformy MTS⁻	720	623

Tabela 6. Wyniki wyszukiwania białek wydłużonych na końcu N (NTE). Kolumna „Cały genom” odnosi się do całego genomu drożdżowego zawartego w SGD. „Geny wysokiego prawdopodobieństwa” powstały przez odfiltrowanie części ORF z całego genomu (patrz Materiały i Metody). Wiersz „Geny” zawiera wszystkie ORF uwzględnione w teście. „Geny NTE” to liczba genów, dla których przewidziano niekanoniczne wydłużenia. „Izoformy NTE” określają ilość izoform generowanych w wyniku niekanonicznej translacji. „Geny bez NTE” opisują ilość genów, dla których nie przewidziano NTE. Wiersze „Białka MTS⁺/MTS⁻” i „Izoformy MTS⁺/MTS⁻” wskazują ilość białek/izoform, którym przypisano parametry MTS⁺ i MTS⁻ (szczegóły w tekście oraz w Materiałach i Metodach).



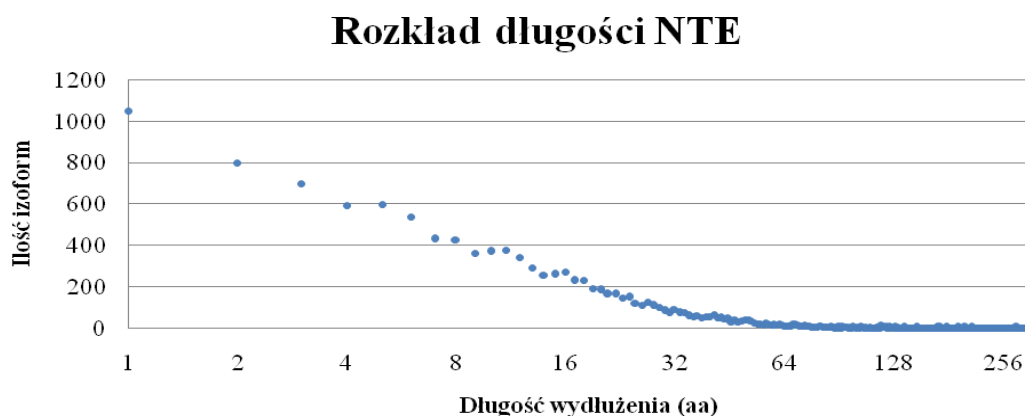
Wykres 1. Udział niekanonicznych miejsc startu ze względu na ilość wykrytych poszczególnych kodonów. Podwójny wykres kołowy. W legendzie wypisano wszystkie kodony niekanoniczne użyte w analizie. uAUG (*upstream AUG*) oznacza dodatkowy kodon AUG położony powyżej kanonicznego miejsca startu. Mniejszy wykres obrazuje ilość kodonów NUG, zaś większy pozostałe, wskazując na udział ogółu kodonów NUG w całości. Wartości procentowe z mniejszego wykresu dotyczą wartości uzyskanych w całej analizie.

Analizując udział kodonów, wykrytych przez algorytm (Wykres 1), można zauważyć przewagę kodonów AUU i AUA, co można tłumaczyć tym, że genom drożdżowy jest AT-bogaty (tylko 38% par GC) (Mewes et al., 1997). Najprawdopodobniej większość przewidzianych aTIS nie jest używana do startu translacji, a zwłaszcza te położone najdalej od kodonu AUG, bądź będące w nieodpowiednim kontekście Kozak (Kozak, 1987). Kodony NUG, najczęściej wykorzystywane w inicjacji translacji (Hinnebusch, et al., 2016; Na et al., 2018), stanowią 29% wszystkich prawdopodobnych miejsc startu.

Ciekawe jest to, że znaleziono 212 przypadków (1,8% wszystkich), dla których kodonem aTIS jest uAUG położony powyżej anotowanego TIS. Może to sugerować błędną anotację genów w bazie SGD. Tak było w sytuacji genów *VASI* (Chatton et al., 1999) i *HTSI* (Chiu et al., 1992), które, podobnie jak *ALAI* i *GRSI*, kodują syntetazy aa-tRNA zarówno cytoplazmatyczne (krótszy ORF) jak i mitochondrialne (wydłużony ORF). W przypadku *ALAI* i *GRSI*, synteza wydłużonej izoformy inicjowana jest z kodonów niekanonicznych, zaś dla *VASI* i *HTSI* alternatywnym miejscem startu są kodony AUG położone powyżej TIS dla formy cytoplazmatycznej. Obecnie w SGD dla tych dwóch genów (*VASI* i *HTSI*) zapisane jest miejsce startu dające najdłuższy CDS. Jednak nie można jednoznacznie przełożyć tych wniosków na pozostałe wykryte przypadki. Poznanie dokładnych miejsc startu inicjacji translacji TIS (and. *translation initiation site*) jest możliwe jedynie dzięki badaniu pojedynczych przypadków oraz analizie profilowania rybosomów z użyciem antybiotyków immobilizujących rybosomy inicjujące, takich jak harringtonina czy laktynomycyna (Andreev

et al., 2017). Obecnie dla *S. cerevisiae* nie są jednak dostępne dane z profilowania rybosomów inicjujących.

Długość NTE waha się od 1 do 299 aminokwasów (aa), co wynika z założeń opisanego w programie algorytmu (analizie poddano 300 aminokwasów powyżej kodonu AUG), natomiast średnia wartość wynosi 18,4 aa, a mediana 10 aa. Przeważają wydłużenia krótkie od 1 do 30 aa (84,1%). Bardzo długie NTE (powyżej 100 aa) są sporadyczne (227 przypadki; 1,9% wszystkich izoform). Współczynnik korelacji Pearsona r wynosi -0,4889, co wskazuje na przewagę krótkich wydłużeń (Wykres 2). Nie wykryto zależności między długością NTE a długością białka (współczynnik korelacji Pearsona $r = -0,0352$), co oznacza, że każde białko może mieć zarówno długie jak i krótkie NTE.



Wykres 2. Rozkład długości wydłużeń na końcu N. Na osi X oznaczono długość izoform (w aminokwasach), zaś na osi Y ilość izoform o danej długości. Wartości na osi X podano w skali logarytmicznej (\log_2).

Najdłuższe przewidziane NTE posiadają ORF *YBL068W-A* i *YBR219C*, jednak jest mało prawdopodobne, że tak długie białka są syntetyzowane, gdyż w transkryptomie drożdżowym mRNA dla tych ORF nie jest wykrywany (Pelechano et al., 2013). Gdyby pominąć izoformy najkrótsze (1-2 aa) oraz najdłuższe (powyżej 100 aa), pozostaje 9 529 izoform, które wykazują jeszcze silniejszą ujemną korelację między długością izoformy a jej częstością (współczynnik korelacji Pearsona $r = -0,7764$), co ponownie wskazuje na znaczną przewagę krótkich NTE. Uzyskane dane współgrają z danymi proteomicznymi końców N, wskazującymi na wykorzystywanie TIS położonych blisko siebie (Helsens, et al., 2011; Van Damme et al., 2014).

Białka NTE przeanalizowano również pod kątem funkcji przypisanych w bazie Gene Ontology (GO). Nie zaobserwowano jednak wysokiego wzbogacenia w którejkolwiek z kategorii (Tabela 7). Nie jest to jednak dziwne, biorąc pod uwagę, że ponad 2/3 białek

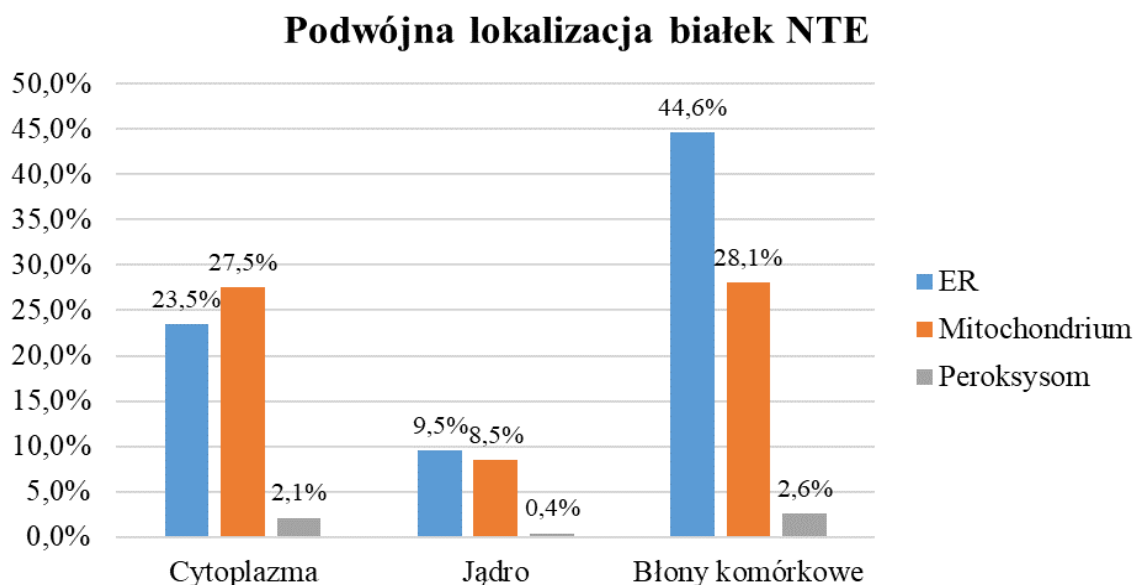
kodowanych w genomie drożdżowym może potencjalnie posiadać przynajmniej jedno wydłużenie na końcu N.

Kategoria Gene Ontology	Występowanie w NTE	Występowanie w genomie	Wartość p
Aktywność katalityczna	39,1%	34,0%	1,33e-22
Wiązanie jonów	26,1%	22,3%	1,37e-16
Wiązanie białek	16,5%	13,6%	6,44e-14
Wiązanie małych cząstek	15,4%	13,2%	1,59e-07
Wiązanie anionów	15,6%	13,5%	7,96e-07
Wiązanie nukleotydów	14,2%	12,2%	1,64e-06
Wiązanie antybiotyków	12,0%	10,2%	2,73e-06
Wiązanie jonów metali	13,4%	11,5%	9,11e-06
Aktywność hydrolazowa	17,0%	15,0%	1,77e-05
Wiązanie ATP	10,9%	9,3%	9,35e-05
Wiązanie rybonukleotydów	12,6%	10,9%	0,00010
Aktywność pirofosfotazowa	7,6%	6,3%	0,00017
Wiązanie mRNA	3,5%	2,7%	0,00022
Aktywność ATPazy	5,2%	4,2%	0,00211
Wiązanie DNA	9,9%	8,6%	0,00424
Wiązanie niezwinionych białek	1,7%	1,2%	0,00476

Tabela 7. Analiza GO Term Function przy użyciu GO Term Finder, dostępnego na stronie SGD. Wybrano szesnaście przykładowych kategorii Gene Ontology. Kolumna „występowanie w NTE” określa ilość białek NTE o danej funkcji w zbiorze białek. Analogicznie „występowanie w genomie” dotyczy tylko częstotliwości danych genów w całym genomie. Część białek posiada kilka funkcji i jest wyszczególnione w kilku kategoriach. Wartość p, wskazująca prawdopodobieństwo losowego dopasowania, została wyliczona według algorytmów oprogramowania.

Przeprowadzono również analizę lokalizacji białek, które mogą potencjalnie zyskiwać NTE. Na początku skupiono się na lokalizacji opisanej w SGD, gdyż jest ona udokumentowana w badaniach wysokoprzepustowych i eksperymentalnych. W szczególności przyjrano się białkom przypisanym do więcej niż jednego organellum, ponieważ w wielu przypadkach białka lokalizują się w kilku miejscach, pomimo braku specyficznych sygnałów kierujących. W kontekście badań nad NTE szczególnie interesująca jest lokalizacja mitochondrialna, peroksysomalna oraz w retikulum endoplazmatycznym (ER), gdyż sygnały kierujące do tych organelli komórkowych mają postać presekwencji na końcu N (Kunze i Berger, 2015). Przyjmując, że białka NTE cechują się podwójną lokalizacją, sprawdzono jaki procent białek cytoplazmatycznych stanowią białka przypisane do ER, mitochondriów i peroksysomów. Takie samo zestawienie przygotowano też dla białek jądrowych i błonowych (Wykres 3). Okazuje się, że ponad 1/4 białek cytoplazmatycznych i błonowych posiada dodatkową lokalizacją mitochondrialną, wykazaną eksperymentalnie, bądź w badaniach wysokoprzepustowych według SGD. Niski udział białek peroksysomalnych w tym zestawieniu wynika z ich niewielkiej ilości w stosunku do całego proteomu drożdżowego.

Warto wspomnieć, że lokalizacja 13,4% białek NTE jest nieznana. Być może dalsze badania nad wydłużeniami końca N pozwolą lepiej poznać ich funkcje oraz lokalizację komórkową.



Wykres 3. Podwójna lokalizacja białek NTE. Wykres kolumnowy, pokazujący udział białek ER (kolor niebieski), mitochondrialnych (kolor pomarańczowy) i peroksysomalnych (kolor szary) we frakcjach białek cytoplazmatycznych, jądrowych oraz błonowych.

Podsumowując, w wyniku inicjacji translacji z kodonów non-AUG mogą powstawać tysiące wariantów NTE o różnych funkcjach. Przeprowadzona analiza jednak nie wyczerpuje wszystkich możliwości startu syntezy białka z kodonów niekanonicznych. Możliwe jest również wykorzystywanie miejsc nie ujętych w algorytmie wyszukiwania wydłużeń. Na przykład, kodon CUU stanowi miejsce startu dla Hyr1 (Kritsiligkou et al., 2017), pomimo że nie jest spokrewniony z AUG.

2. Identyfikacja MTS w izoformach wydłużonych na końcu N

Aby poznać lepiej funkcje białek NTE, postanowiono przyjrzeć się dokładniej lokalizacji mitochondrialnej poszczególnych izoform. Wybrano mitochondria ze względu na dostępność wielu wiarygodnych algorytmów przewidujących wystąpienie sygnału lokalizacji mitochondrialnej MTS oraz istnienie dobrze opracowanych metod izolacji mitochondriów, co jest istotne dla eksperymentalnej weryfikacji danych. Ponadto, wcześniejsze doniesienia wskazują kilka przypadków mitochondrialnej lokalizacji wydłużonych na końcu N izoform białkowych syntetyzowanych z kodonu non-AUG (Chang i Wang, 2004; Tang et al., 2004; Soumi et al., 2014; Kritsiligkou et al., 2017). Postanowiono więc sprawdzić, czy takie zjawisko jest powszechne i znaczące dla różnorodności drożdżowego genomu mitochondrialnego.

Prawdopodobieństwo wystąpienia MTS wyliczono przy użyciu programów MitoProt II (Claros et al., 1995), TargetP (Emanuelsson et al., 2007) oraz MitoFates (Fukusawa et al., 2015) dla form kanonicznych oraz każdej izoformy wydłużonej. Na podstawie uzyskanych wyników nadano parametry MTS^+ oraz MTS^- (patrz Materiały i Metody), które określają w jaki sposób wydłużone izoformy zmieniają swoją lokalizację mitochondrialną. Białka MTS^+ to takie, których forma podstawowa jest niemitochondrialna, a wydłużona lokalizuje się w mitochondriach, natomiast parametr MTS^- oznacza sytuację odwrotną, kiedy to wydłużone formy pierwotnie mitochondrialnych białek tracą lokalizację mitochondrialną.

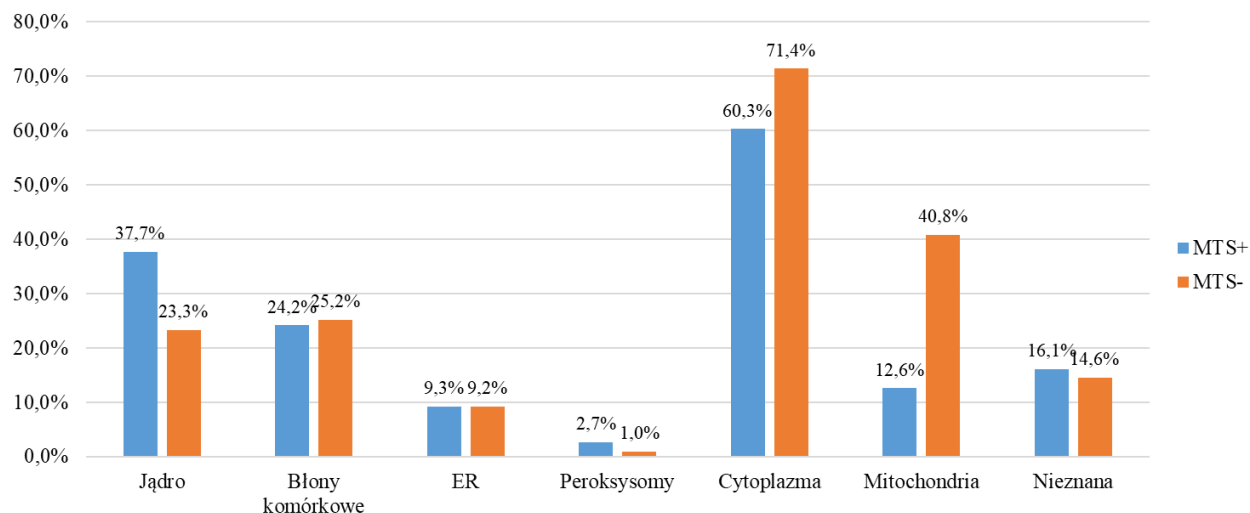
Długość natywnych MTS u drożdży *S. cerevisiae* waha się od 6 do 122 aminokwasów (Vögtle et al., 2009; Teixeira i Glaser, 2013), dlatego postanowiono sprawdzić zależność między długością NTE a prawdopodobieństwem uzyskania MTS przez izoformę. Nie wykryto jednak żadnej znaczącej korelacji dla wszystkich izoform NTE (współczynnik korelacji Pearsona $r = 0,014$), ani dla białek MTS^+ ($r = 0,196$) czy dla MTS^- ($r = -0,258$), dlatego skupiono się na innych cechach zbiorów MTS^+ i MTS^- .

Parametr MTS^+ został przypisany 517 białkom, dającym w sumie 1 199 izoform. Wśród nich znaleziono Ala1, Grs1, Hfa1 i Hyr1, dla których wcześniej opisano wydłużone izoformy lokalizujące się w mitochondriach (Chang i Wang, 2004; Tang et al., 2004; Soumi et al., 2014; Kritsiligkou et al., 2017). Tendencja ogólnej lokalizacji komórkowej białek MTS^+ jest podobna do białek MTS^- , lecz można zaobserwować znacznie wyższy udział białek jądrowych (Wykres 4a i 4b). Większość białek jądrowych odpowiada za obróbkę DNA

i RNA oraz ich szeroko pojęty metabolizm. Mitochondria również posiadają własny genom, stąd białka jądrowe w mitochondriach mogą pełnić analogiczne funkcje.

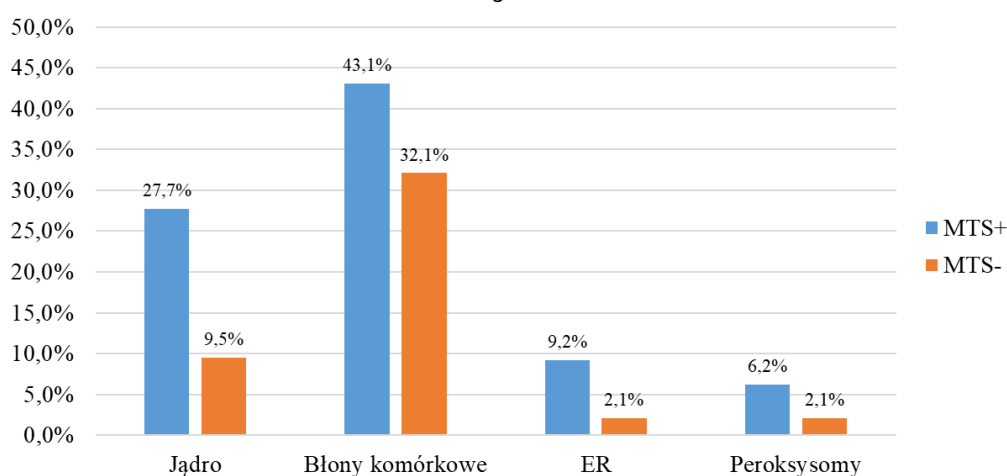
A

Lokalizacja białek MTS⁺ i MTS⁻



B

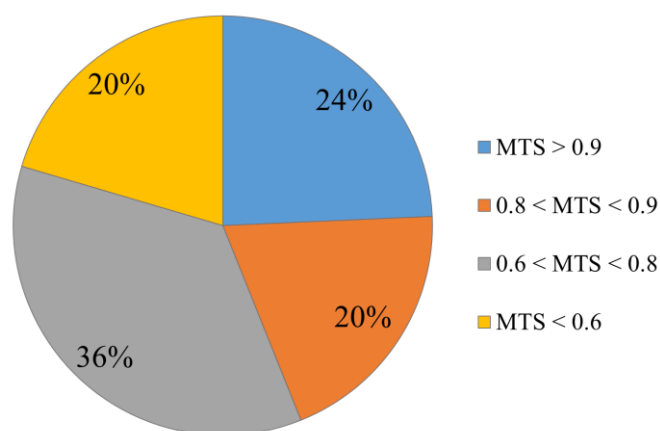
Dodatkowa lokalizacja białek MTS⁺ i MTS⁻



Wykres 4. Analiza lokalizacji białek MTS⁺ i MTS⁻. Zestawiono lokalizację komórkową białek kategorii MTS⁺ (kolor niebieski) i MTS⁻ (kolor pomarańczowy) według SGD. A) Wybrane lokalizacje komórkowe wszystkich białek o parametrach MTS⁺ i MTS⁻. Ponieważ białka cechują się lokalizacją w kilku miejscach (oprócz lokalizacji nieznannej), wartości procentowe nie sumują się do 100%. B) Wybrane dodatkowe lokalizacje białek MTS⁺ i MTS⁻, anotowanych jako mitochondrialne w SGD. Wykluczono białka cytoplazmatyczne, gdyż w SGD każde białko mitochondrialne jest jednocześnie przydzielone do puli białek lokalizujących się w cytoplazmie.

Ciekawe jest, że analiza GO Term Function dla białek MTS⁺ nie wykryła żadnego wzbogacenia jakiegokolwiek frakcji białek, co wskazuje na to, że MTS w obrębie NTE mogą zyskiwać całkowicie losowe białka. Wiele z białek MTS⁺ zyskuje bardzo wysokie parametry MTS (543 izoform z MTS > 0,8 oraz 301 z MTS > 0,9) (Wykres 5), dzięki czemu

prawdopodobieństwo ich lokalizacji mitochondrialnej znacząco wzrasta. Ponadto, kilkanaście z nich (6,2%) z bardzo niskim, lecz rozpoznawalnym MTS w kanonicznej formie (poniżej 0,2), zyskuje bardzo wysoki MTS w formie wydłużonej (ponad 0,9).



Wykres 5. Prawdopodobieństwo uzyskania MTS przez białka MTS⁺. Liczby oznaczają maksymalną wartość MTS, wyliczoną przez MitoProt II, TargetP i MitoFates.

Z 517 białek, oznaczonych parametrem MTS⁺, tylko 88 jest opisanych w SGD jako mitochondrialne, ale 364 zidentyfikowano w najnowszym opracowaniu proteomu mitochondrialnego, z czego 59 zakwalifikowano do najwyższej, pierwszej klasy czystości (Morgenstern et al., 2017). Interesujące jest to, że tylko 34% białek określonych jako mitochondrialne ma przewidziany MTS z prawdopodobieństwem powyżej 0,4 (według MitoProt II, TargetP lub MitoFates). W związku z tym sprawdzono, ile z pozostałych mitochondrialnych białek o niskim prawdopodobieństwie lokalizacji mitochondrialnej zyskuje MTS poprzez niekanoniczne wydłużenia na końcu N. Okazało się, że więcej niż 14% białek NTE i zidentyfikowanych w badaniu proteomicznym jako mitochondrialne należy do kategorii MTS⁺ (319 białek/675 izoform). Sugeruje to, że obecność MTS na NTE wpływa na wzbogacenie proteomu mitochondrialnego, chociaż nie wyjaśnia w pełni fenomenu bogactwa mitoproteomu.

Potwierdzeniem mitochondrialnej lokalizacji izoform MTS⁺ może być ich funkcja mitochondrialna i wpływ na funkcjonowanie mitochondriów. W związku z tym sprawdzono najpierw, czy brak białek MTS⁺ wywołuje fenotyp defektu oddechowego *petite* (Slonimski i Ephrussi, 1949; Chen i Clark-Walker, 2000; Dimmer et al., 2002; Merz i Westermann, 2009). Spośród 554 genów *pet*, których delecja powoduje fenotyp *petite*, cechujący się brakiem wzrostu na niefermentowalnym źródle węgla, 27 należało do kategorii MTS⁺, a 355 ogółem do białek NTE (z czego 38 to białka MTS⁻). Warto jednak wspomnieć, że klasyfikacja genów *pet* nie obejmuje białek niezbędnych do przeżycia komórki. Sprawdzone więc,

ile z białek kluczowych dla funkcjonowania drożdży, których deplecja dzięki ekspresji pod kontrolą regulowanego przez obecność tetracykliny promotora *Tet* powoduje defekty mitochondrialne (Altmann i Westermann, 2005), należy do kategorii MTS⁺. Spośród 119 takich białek 9 należało do kategorii MTS⁺, 85 ogółem do białek NTE, a 3 do puli białek MTS⁻ (Tabela 8).

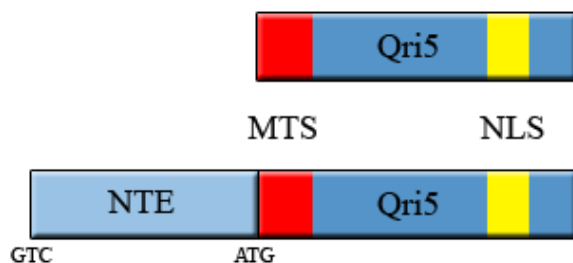
Rodzaj badania	Ogół białek	Białka NTE	Białka MTS ⁺	Białka MTS ⁻
Geny, których delecja powoduje fenotyp <i>petite</i> (Dimmer et al., 2002; Merz i Westermann, 2009; SGD)	554	355	27	38
Geny niezbędne do przeżycia komórki, których wyciszenie powoduje defekty oddechowe (Altmann i Westermann, 2005)	119	85	9	3
Razem	673	440	36	41

Tabela 8. Zestawienie udziału białek NTE, MTS⁺ i MTS⁻ w puli białek niezbędnych do funkcjonowania mitochondriów. Kolumny odnoszą się odpowiednio do rodzaju badania, ogółu białek ujętych w badaniu oraz poszczególnych kategorii (NTE, MTS⁺ i MTS⁻). Wiersze tyczą się odpowiednich badań wraz z referencjami (szczegóły w tekście). We wierszu „Razem” podano wartości sumaryczne dla obu rodzajów analiz.

Parametr MTS⁻ przypisano 206 białkom, dającym 623 izoform, potencjalnie zyskującym nową lokalizację. Wiele z nich to białka związane z metabolizmem mitochondriów i oddychaniem komórkowym (84 z nich jest określone jako mitochondrialne w SGD), stąd ich funkcje poza mitochondriami są zagadkowe. Większość z białek MTS⁻ według bazy SGD lokalizuje się w cytoplazmie, co spowodowane jest tym, że każde białko mitochondrialne jest automatycznie uznane jako cytoplazmatyczne, dlatego zbiór ten może być zawierać wiele fałszywie pozytywnych wyników. Białka o lokalizacji nieznanej stanowią około 15%, co jednak nie odbiega znacząco od udziału tej frakcji w całej populacji białek NTE (13,4%) (Wykres 4a). Kiedy sprawdzono jaką dodatkową pozamitochondrialną lokalizację posiadają białka MTS⁻, zauważono wyraźny udział białek błonowych (ponad 30%; Wykres 4b). Jednak ta obserwacja również może być nie do końca poprawna, gdyż wiele białek wewnętrznej i zewnętrznej błony mitochondrialnej, które uzyskały parametr MTS⁻, zalicza się do ogólnej puli białek błonowych. Dla wielu izoform MTS⁻ utrata MTS w stosunku do formy podstawowej jest znacząca, a w przypadku aż 29,9% wyjątkowo wysoka (ponad 80%). Trudno jednak stwierdzić bez potwierdzenia eksperymentalnego, czy ta pula białek faktycznie ma zmienioną lokalizację, ponieważ transportery mitochondrialne mogą również rozpoznawć wewnętrzne MTS (ang. *internal MTS like*; iMTS-Ls) (Backes et al., 2018).

Co ciekawe, 8 białek posiada dwa typy izoform: jedne zaklasyfikowane jako MTS⁺, a drugie jako MTS⁻.

Interesującym przykładem białka MTS⁻ jest białko wewnętrznej błony mitochondrialnej Qri5, które jest wymagane do akumulacji dojrzałego mRNA *COX1* i może brać udział w translacji białka Cox1. Poza silnym MTS na końcu N, Qri5 posiada sygnał lokalizacji jądrowej NLS (ang. *nuclear localisation signal*) na końcu C (wyliczony programami PSORT, NucPred i cNLS Mapper; patrz Materiały i Metody). W przypadku wydłużonej formy, MTS na końcu N jest zakrywany przez 28-aminokwasową sekwencję, przez co według predykcji, NTE-Qri5 lokalizuje się w jądrze komórkowym (Fig 11). Co prawda, według SGD Qri5 jest białkiem błonowym, zlokalizowanym w mitochondriach, ale jest możliwe, że alternatywna forma Qri5 jest obecna i pełni rolę w jądrze.



NLS według			MTS według		
PSORT	NucPred	cNLS Mapper	MitoProt II	TargetP	MitoFates
82,6%	0,86	12	0,984	0,840	0,950
91,3%	0,92	12	0,055	0,192	0,018

Fig 11. Qri5 – przykład białka MTS⁻. Górny panel przedstawia schemat białka Qri5. Na czerwono zaznaczono sygnał lokalizacji mitochondrialnej MTS, a na żółto sygnał lokalizacji jądrowej NLS. Kolorem jasnoniebieskim wskazano NTE, a kodony inicjujące zapisano poniżej. W tabeli przedstawiono prawdopodobieństwo występowania NLS i MTS oszacowane przy użyciu programów PSORT, NucPred, cNLS Mapper (NLS) oraz MitoProt, TargetP i MitoFates (MTS). W przypadku cNLS Mapper podano najwyższą wartość wykrytego sygnału NLS.

Aby dokładniej poznać dokładniej funkcje białek MTS⁻, należy najpierw przeprowadzić bardziej szczegółowe analizy, łącznie wyliczeniem prawdopodobieństwa występowania w innych organellach komórkowych, a następnie eksperymentalnie potwierdzić ich obecność i działanie w odpowiednich lokalizacjach.

Podsumowując, według wyliczeń bioinformatycznych NTE oraz obecny na nim MTS mają istotny wpływ na transport białek do mitochondriów. Białka NTE pełnią funkcje w mitochondriach, wpływając na ich fenotyp i funkcje w komórce. Przeprowadzone analizy bioinformatyczne stanowią podstawę dla kolejnych badań eksperymentalnych.

3. Lokalizacja białek NTE w mitochondriach

Aby potwierdzić eksperymentalnie lokalizację mitochondrialną białek NTE (MT-NTE; mitochondrialnych białek NTE), wybrano 18 losowych białek z kategorii MTS⁺ (Tabela 9) i przeprowadzono frakcjonowanie komórek drożdżowych szczepów, wyrażających analizowane białka w fuzji ze znacznikiem TAP-tag (ang. *Tandem Affinity Purification tag*) na końcu C, po czym za pomocą techniki western blot oznaczono obecność badanych białek w poszczególnych frakcjach. Cztery z wybranych białek (Adh4, Trr1, Trz1 i Hyr1) zostały już wcześniej opisane jako mitochondrialne (Young i Pilgrim, 1985; Pedrajas et al., 1999; Trotter i Grant, 2005; Vögtle et al., 2012; Skowronek et al., 2014; Kritsiligkou et al., 2017), co potwierdziły uzyskane w pracy wyniki.

Białko	Kanoniczny MTS (średnia MP/TP/MF)	NTE MTS (średnia MP/TP/MF)	Długość NTE	Klasa (Morgenstern et al., 2017)
Adh4	0,032	0,899	3, 16, 38, 43, 56 , 69, 70, 79, 80, 83, 86, 89	Klasa 1
Atg1	0,049	0,273	3 , 15	Klasa 1
Fur1	0,089	0,292	3, 18, 24, 29, 35	Klasa 3
Gpp1	0,036	0,778	2, 5, 12, 21, 30 , 32, 33, 39	Klasa 2
Gpp2	0,039	0,487	3, 6, 12, 17, 24, 31	Klasa 3
Hyr1	0,023	0,452	1, 10, 16, 24, 26 , 30	Klasa 3
Ioc4	0,040	0,442	1, 7, 14, 17, 18	-
Lsm1	0,082	0,291	5 , 21	Klasa 3
Lsm4	0,042	0,190	6, 11 , 17, 18	-
Mtr4	0,042	0,156	3 , 6	Klasa 3
Mum2	0,050	0,317	12, 17	-
Nop9	0,090	0,319	1, 3, 4, 12, 13 , 14	Klasa 1
Pmt2	0,144	0,596	1, 23, 36, 43, 45 , 65, 70, 75, 76, 78, 82, 85, 92	Klasa 3
Rnt1	0,034	0,338	7, 26, 27, 33, 39, 43 , 48, 64, 69, 72	-
Trm3	0,107	0,289	12	Klasa 3
Trr1	0,108	0,363	1, 4, 5, 9, 13, 16, 21, 28, 38	Klasa 3
Trx2	0,070	0,565	1, 8, 9, 23, 24, 30, 31	Klasa 3
Trz1	0,066	0,866	1, 14, 33, 43, 51 , 54	Klasa 1
Sgt2	0,136	0,150	1, 4, 6, 9	Klasa 3

Tabela 9. Lista przebadanych białek z kategorii MTS⁺. Kolumny „Kanoniczny MTS” i „NTE MTS” zawierają przewidzianą wartość MTS (podano średnią z parametrów wyliczonych przez MitoProt II (MP), TargetP (TP) i MitoFates (MF)) odpowiednio dla formy syntetyzowanej z kodonu AUG i non-AUG. W przypadku „NTE MTS” przedstawione są wartości najwyższe. „Długość NTE” określa długość potencjalnych wydłużeń na końcu N (NTE), powstałych w wyniku inicjacji translacji z kodonu non-AUG. Pogrubiono izoformę o najwyższej średniej wartości MTS. „Klasa” odnosi się do klasyfikacji białek według najnowszego proteomu mitochondrialnego (Morgenstern et al., 2017).

Frakcjonowaniu poddano ekstrakty komórkowe ze szczepów drożdżowych, posiadających w swoim genomie badane białka w fuzji ze znacznikiem TAP-tag na końcu C przy zachowaniu natywnego promotora i końca N. Wykorzystano zmodyfikowany protokół

do izolacji mitochondriów z drożdży (Meisinger et al., 2006; Meisinger et al., 2000), umożliwiając częściowo ilościowe uzyskanie poszczególnych frakcji i ich późniejsze porównanie (patrz Materiały i Metody).

Białka znakowane na końcu C znacznikiem TAP-tag wykrywano za pomocą przeciwciał PAP (ang. *peroxidase-anti-peroxidase*), skierowanych przeciwko znacznikowi. W przypadku Trr1 użyto specyficznych poliklonalnych przeciwciał. W celu określenia czystości poszczególnych frakcji zastosowano specyficzne przeciwciała przeciwko mitochondrialnemu Ilv2 i cytoplazmatycznemu Sgt2. Dodatkowo, Sgt2, będące białkiem NTE bez wykrytego MTS, służyło za kontrolę negatywną metody frakcjonowania (Fig 12-13).

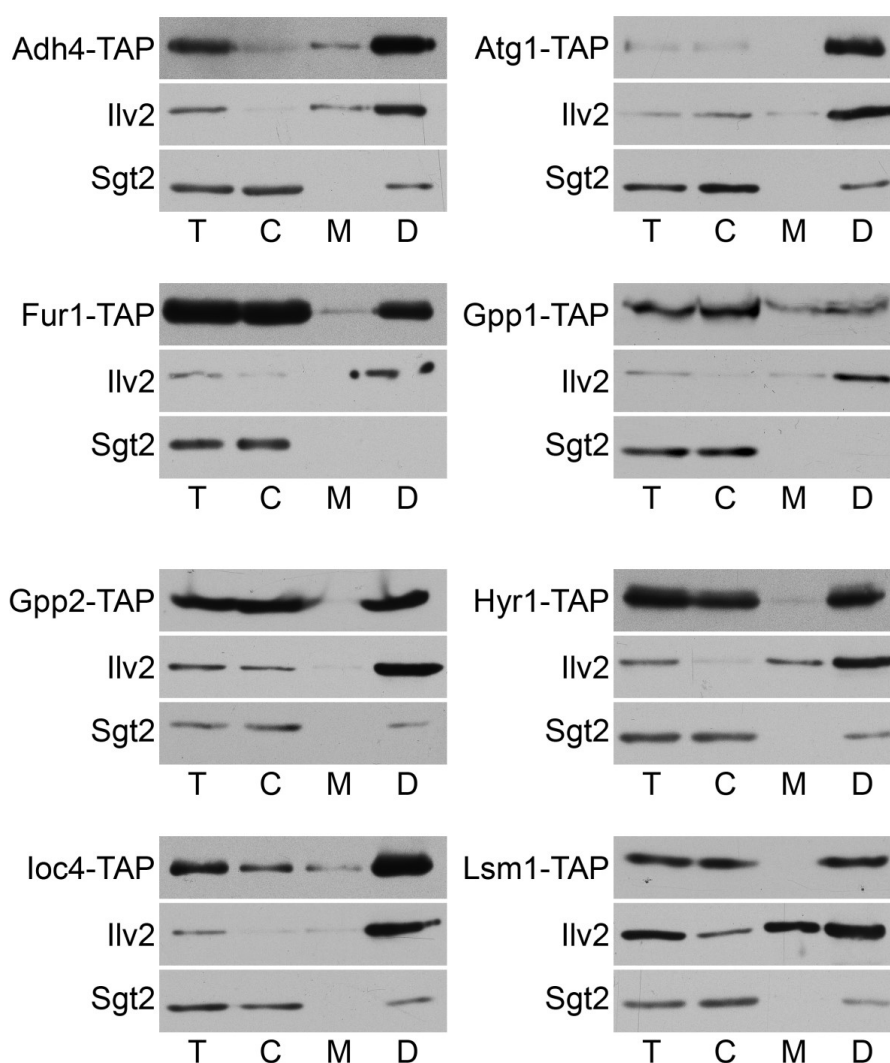


Fig 12. Analiza western blot ekstraktów drożdżowych białek MTS⁺ po frakcjonowaniu. Przedstawiono ekstrakt całkowity (frakcja T), frakcję cytozolową (C), mitochondrialną (M) oraz mitochondrialną zagęszczoną (D; 8 razy skoncentrowana). Białka znakowane na końcu C znacznikiem TAP-tag wykryto przy użyciu przeciwciał PAP; Ilv2 (kontrola czystości frakcji mitochondrialnej) oraz Sgt2 (kontrola czystości frakcji cytoplazmatycznej) wykryto za pomocą specyficznych przeciwciał.

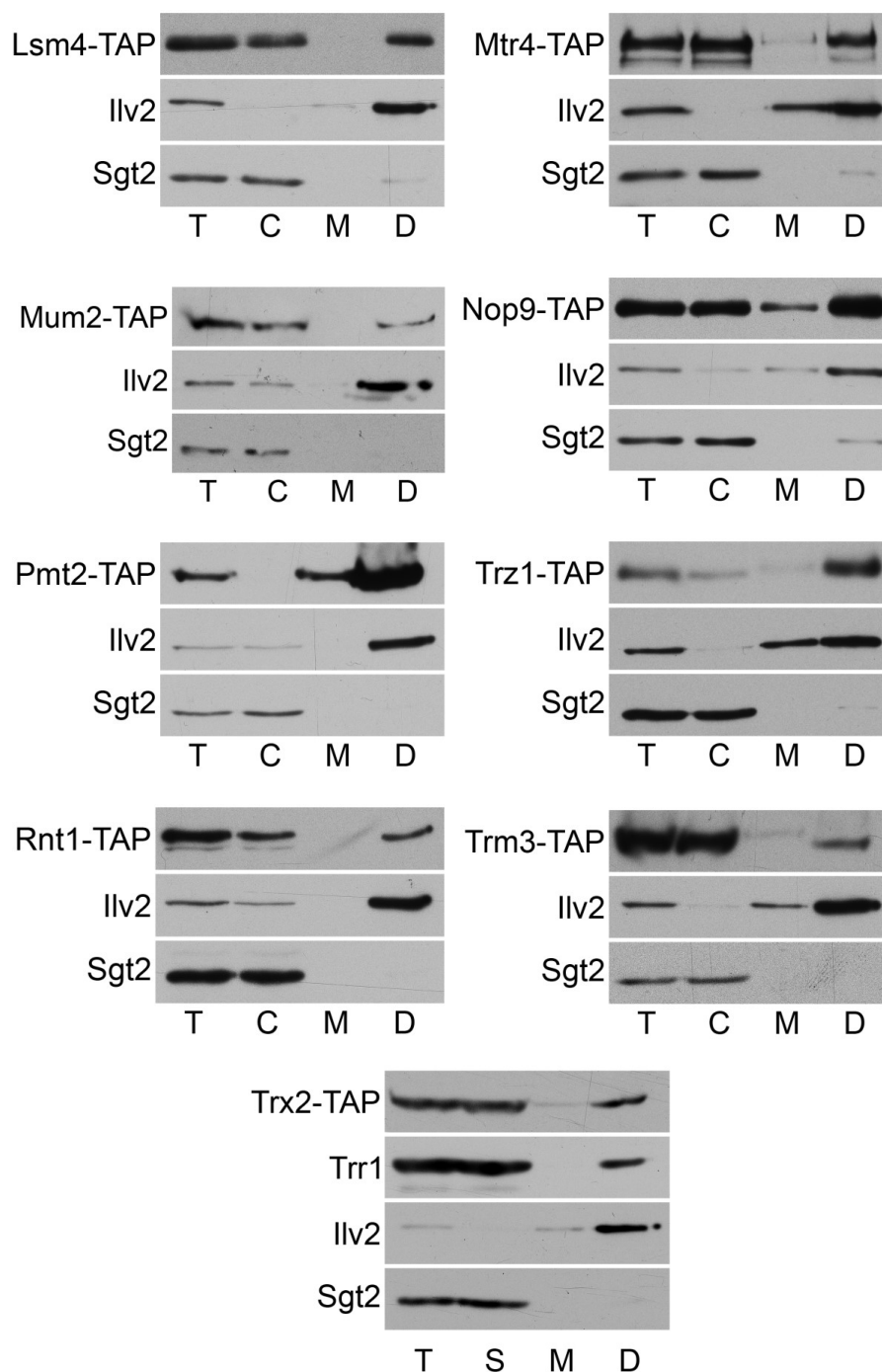


Fig 13. Analiza western blot ekstraktów drożdżowych białek MTS⁺ po frakcjonowaniu, ciąg dalszy. Opis jak przy Fig 12. Trr1 wykryto za pomocą specyficznych przeciwciał.

Inicjacja startu translacji z kodonu niekanonicznego zachodzi przeważnie na bardzo niskim poziomie, stąd nie zawsze możliwe było zaobserwowanie sygnału w podstawowej frakcji mitochondrialnej (M). Aby umożliwić detekcję niewielkiej ilości białek, przygotowano również frakcję zagęszczonych mitochondriów D (8 razy bardziej skoncentrowana niż frakcja

M; szczegóły w Materiałach i Metodach), w której można zaobserwować wyraźny sygnał pochodzący od badanego białka.

Wszystkie z testowanych białek MTS⁺ wykryto we frakcjach mitochondrialnych, głównie we frakcji zagęszczonej, co jest zgodne z spodziewanym niskim poziomem wydłużonych izoform. Kontrolne cytoplazmatyczne białko Sgt2 nie było obecne we frakcjach M i D (Fig 12-13). Analizy te wskazują na to, że podwójna lokalizacja białek, cytozolowa i mitochondrialna, jest powszechniejsza niż przewidywano i może mieć związek z występowaniem dodatkowego MTS w obrębie NTE. Jednak na tym etapie nie można wykluczyć możliwości, że badane białka lokalizują się w mitochondriach nie jako izoformy NTE, ale jako formy podstawowe transportowane do mitochondriów w niekanoniczny sposób. Ponadto, w większości przypadków białek mitochondrialnych presekwencja zawierająca MTS jest proteolitycznie odcinana, co utrudnia stwierdzenie obecności regionu NTE dla białek MT-NTE.

Aby potwierdzić, że do mitochondriów kierowane są wydłużone formy białek, a nie kanoniczne, przeprowadzono frakcjonowanie ekstraktów komórkowych ze szczepów drożdżowych, wyrażających białka tylko w kanonicznej formie, inicjowanej z kodonu AUG. Do tego celu wybrano białka Lsm1 i Mtr4 wyrażane w fuzji ze znacznikiem HA (ang. *human influenza hemagglutinin*) (Fig 14). W przypadku Mtr4 użyto szczepu GAL1::HA-Mtr4^{AUG} (Torchet et al., 2002), zaś dla Lsm1 wykorzystano szczep wyrażający GAL1-Lsm1^{AUG}-HA z plazmidu kolekcji MORF (ang. *movable ORF*) (Gelperin et al., 2005). Szczep GAL1-Lsm1^{AUG}-HA hodowano na pożywce YPD, by uniknąć błędnej lokalizacji, wynikającej z nadekspresji, zaś GAL1::HA-Mtr4 na YPGal, gdyż deplecja Mtr4 jest dla drożdży letalna.

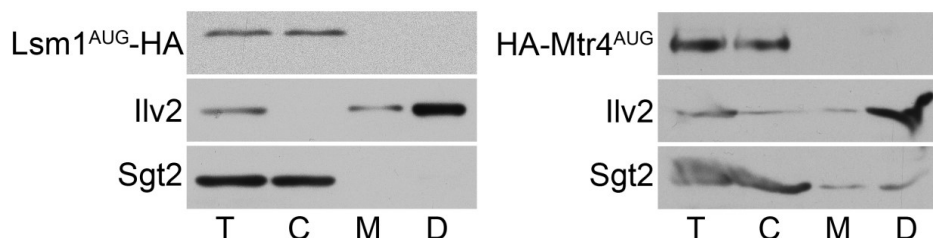


Fig 14. Frakcjonowanie szczepów, wyrażających jedynie kanoniczne formy Mtr4 i Lsm1. Analiza western blot ekstraktu całkowitego (T), frakcji cytozolowej (C) i frakcji mitochondrialnych (M, D). Białka znakowane znacznikiem HA zostały wykryte przy użyciu przeciwciał α -HA. Ilv2 i Sgt2 jak w figurach 12-13.

Tak jak przewidywano, formy podstawowe tych białek nie były obecne w mitochondriach, co potwierdza koncepcję o istotnym udziale NTE i zawartego na nim MTS w lokalizacji mitochondrialnej białek niemitochondrialnych.

Ciekawą obserwacją, która również popiera tę tezę jest to, że w dwóch przypadkach zbadanych białek MT-NTE, a mianowicie dla Trz1 i Mtr1, obserwuje się nieznaczną, lecz zauważalną, różnicę w migracji między formami cytozolowymi a mitochondrialnymi. Możliwe, że warianty mitochondrialne ulegają modyfikacjom post-translacyjnym w mitochondriach, co powoduje ich wolniejszą migrację (Stram i Payne, 2016), lecz nie można wykluczyć, że ich presekwencje zawierające MTS nie ulegają cięciu proteolitycznemu w mitochondriach, co pozwala na zaobserwowanie obu izoform.

Przebadano również 5 białek spoza puli MTS⁺, o słabym MTS w formie wydłużonej (Tabela 10). Tak jak poprzednio, poddano frakcjonowaniu szczepy drożdżowe, wyrażające badane białka wyznakowane na końcu C znacznikiem TAP-tag. Zaskakujące jest to, że wszystkie z nich lokalizowały się w zagęszczonych frakcjach mitochondrialnych, a cztery z nich nawet we frakcji niezagęszczonej M (Fig 15).

Białko	Kanoniczny MTS (średnia MP/TP/MF)	NTE MTS (średnia MP/TP/MF)	Długość NTE	Klasa (Morgenstern <i>et al.</i> , 2017)
Air2	0,046	0,085	13, 17, 23	-
Naf1	0,081	0,082	1, 3	-
Nrd1	0,059	0,151	1, 26	Klasa 3
Rna14	0,030	0,040	7	Klasa 3
Rna15	0,258	0,235	9, 20, 22, 28, 32, 33	-

Tabela 10. Lista przebadanych białek z kategorii nienależących do kategorii MTS⁺. Opis analogicznie do Tabeli 9.

Wyniki te wskazują na to, że mechanizm transportu do mitochondriów nie jest tak oczywisty i że istnieją ścieżki kierowania białek do mitochondriów niezależne od obecności MTS w badanym białku. We wszystkich tych przypadkach nie można wykluczyć błędów w obliczaniu MTS przez dostępne oprogramowanie, lecz możliwe są też inne przyczyny takiego zjawiska, takie jak transport w wyniku oddziaływania z białkami posiadającymi MTS w formie podstawowej albo NTE (patrz Dyskusja).

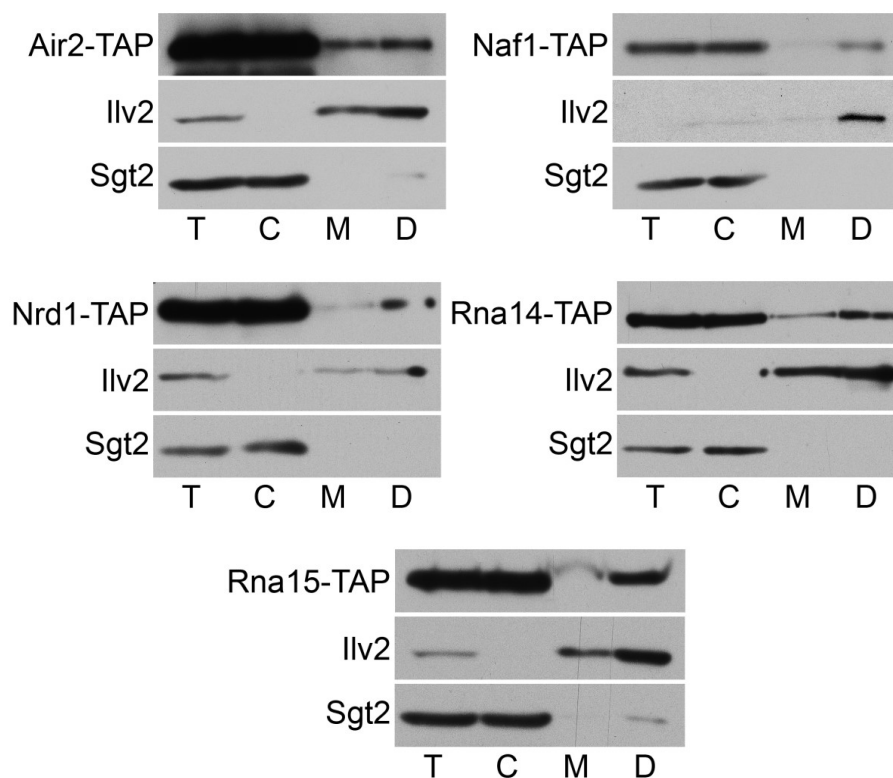


Fig 15. Analiza western blot ekstraktów drożdżowych białek spoza puli MTS⁺. Opis jak przy Fig 12.

Podsumowując, potwierdzono lokalizację mitochondrialną dla 18 białek MT-NTE, które należą do kategorii MTS⁺ i są kierowane do mitochondriów prawdopodobnie dzięki MTS w obrębie NTE, co zostało wykazane dla 2 białek. Ponadto, ujawniono obecność w mitochondriach dodatkowych 5 białek NTE, które w formie wydłużonej nie zyskują MTS i być może są kierowane do mitochondriów dzięki oddziaływaniu z białkami posiadającymi MTS, albo w formie podstawowej albo NTE. Tylko 5 z nich (Adh4, Trr1, Trz1, Hyr1 i Rna14) jest opisanych jako mitochondrialne w SGD, ale 17 z nich zidentyfikowano w najnowszym opracowaniu proteomu mitochondrialnego, przy czym 4 zaklasyfikowano do najwyższej pierwszej klasy czystości (Morgenstern et al., 2017).

4. Potwierdzenie funkcji mitochondrialnych białek MT-NTE

Wiadomo jest, że brak niektórych białek mitochondrialnych może prowadzić do defektów oddechowych, objawiających się m.in. fenotypem *petite* (Dimmer et al., 2002; Merz i Westermann, 2009) i niemożnością wzrostu na pożywce zawierającej niefermentowalne źródło węgla (np. glicerol) (Chen i Clark-Walker, 2000). Postanowiono sprawdzić, czy brak wybranych białek MT-NTE powoduje spowolniony bądź ograniczony wzrost na pożywce YPG. Spośród 10 przetestowanych szczepów posiadających delecje odpowiednich genów, tylko dwa, *lsm1Δ* i *lsm6Δ*, wykazywały wyraźny deficyt oddechowy, objawiający się znacznie spowolnionym wzrostem na pożywce zawierającej glicerol jako jedyne źródło węgla (Fig 16).

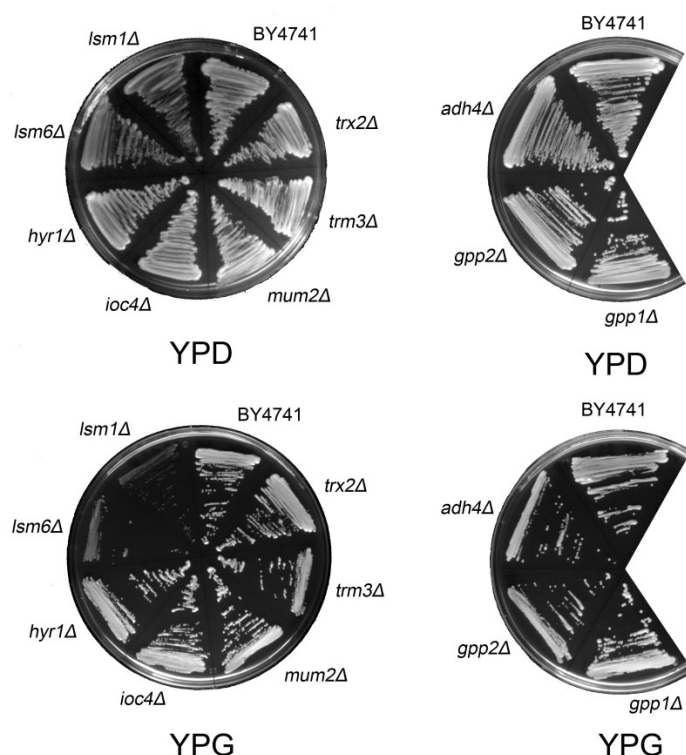


Fig 16. Testy wzrostowe szczepów nie wyrażających białek MT-NTE. Brak Lsm1 i Lsm6 prowadzi do fenotypu deficytu oddechowego. Górny panel przedstawia wzrost na podłożu pełnym z glukozą (YPD), a dolny na podłożu pełnym z glicerolem (YPG). Szczep dziki (BY4741) służy jako referencja.

Ponadto, fenotypem *petite* cechują się szczepy wyrażające tylko formę podstawową, inicjowaną z kodonu AUG białek Mtr4 i Trz1, które są niezbędne do życia dla drożdży (Fig 17; Skowronek et al., 2014). Obserwacje dla Trz1 i Lsm1 są zgodne z dotychczasowymi doniesieniami, wskazującymi że brak tych białek prowadzi do defektów oddechowych (Dimmer et al., 2002; Merz i Westermann, 2009; Skowronek et al., 2014). Zaburzenia w funkcjonowaniu mitochondriów są także obserwowane w przypadku braku Atg1, białka związanego z autofagią (Zhang et al., 2007).

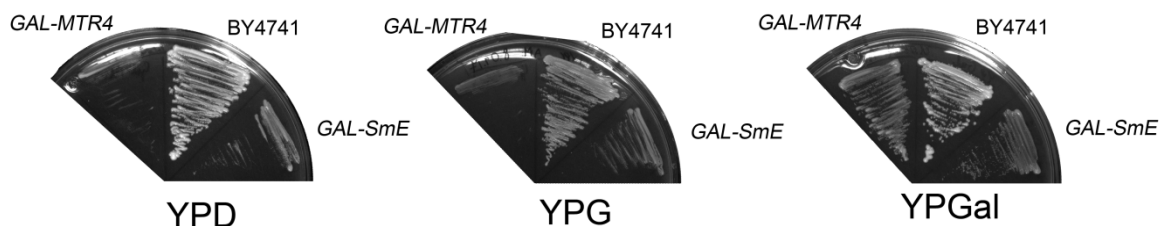


Fig 17. Testy wzrostowe szczepu GAL-MTR4. Szczep GAL-MTR4 wykazuje fenotyp defektu oddechowego. Szczepy: dziki (BY4741), GAL-MTR4 i GAL-SmE hodowano na podłożach pełnych z glukozą (YPD), glicerolem (YPG) oraz galaktozą (YPGal). Szczep GAL-SmE użyto jako kontrolę pozytywną (bez defektu oddechowego) dla GAL-MTR4. Szczepy GAL-SmE i GAL-MTR4 nie rosną przy braku galaktozy (na szalkach YPD) z powodu deplecji niezbędnych do życia białek SmE i Mtr4.

Dokładniejszym analizom poddano endonukleazę 3' tRNA Trz1, która pełni swoje funkcje zarówno w jądrze jak i w mitochondriach (Chen et al., 2005; Skowronek et al., 2014). Aby potwierdzić, że tylko izoforma NTE-Trz1 jest odpowiedzialna za funkcje mitochondrialne tego białka, przygotowano szereg szczepów *trz1Δ*, wyrażających na plazmidach różne warianty Trz1 zgodnie ze schematem (Fig 18). Jest to konieczne, gdyż delecja genu *TRZ1* jest dla drożdży letalna. W tym celu przygotowano następujące konstrukty: *TRZ1^{ATG}* (ekspresja białka Trz1^{AUG} z kanonicznego kodonu AUG), *TRZ1^{TTC}* i *TRZ1^{TTC}-TAP* (ekspresja Trz1^{UUC} lub Trz1^{UUC}-TAP z najbardziej odległego przewidzianego kodonu non-AUG wraz z ekspresją Trz1^{AUG} lub Trz1^{AUG}-TAP z kodonu AUG) oraz *TRZ1^{TTC/GCG}-TAP* z kodonem AUG zmutowanym na GCG (ekspresja wyłącznie Trz1^{UUC}-TAP). Wybrano kodon GCG ponieważ, jako niespokrewniony z AUG, jest

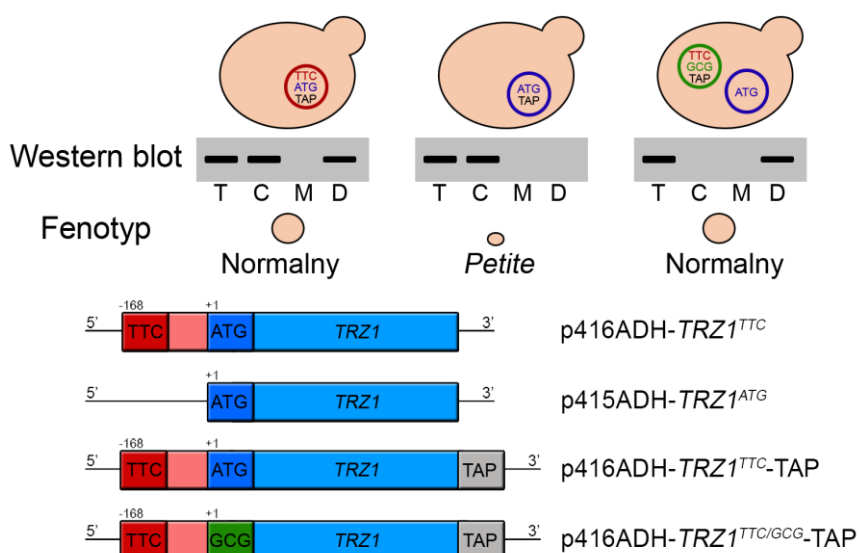


Fig 18. Schemat doświadczenia potwierdzającego lokalizację mitochondrialną NTE-Trz1. Na górnym panelu przedstawiono szczepy wyrażające różne warianty Trz1 i ich spodziewaną lokalizację oraz fenotyp. Poniżej zaprezentowano schemat konstruktyw. TTC i ATG wskazują odpowiednio niekanoniczne i kanoniczne kodony startu translacji; TAP – znacznik TAP-tag; p41XADH – wektory, użyte do konstrukcji plazmidów z wariantami *TRZ1*. Dokładny opis otrzymania konstruktyw w Materiałach i Metodach.

niezwykle rzadko używany jako aTIS (Chang et al., 2010). Jako, że jądrowa izoforma Trz1 jest niezbędna do życia, szczep wyrażający konstrukt *TRZ1^{TTC/GCG}-TAP* posiadał jednocześnie plazmid kodujący kanoniczną izoformę Trz1^{AUG} bez znacznika. Wszystkie szczepy uzyskano w wyniku rozkładu tetrad szczepów diploidalnych *TRZ1/trz1Δ* transformowanych odpowiednimi plazmidami.

Tak jak przewidywano, analiza fenotypu opisanych szczepów wykazała, że obecność obu wariantów Trz1 w komórce, Trz1^{AUG} i Trz1^{UUC}, pozwala na wydajne oddychanie i normalny wzrost na podłożu zawierającym glicerol. Dla komórek wyrażających tylko izoformę Trz1^{AUG} zaobserwowano fenotyp *petite* i brak wzrostu na podłożu z glicerolem (Fig 19). Ponadto, jeżeli komórki posiadające oba konstrukty, *TRZ1^{TTC}* oraz *TRZ1^{ATG}*, utracą wariant *TRZ1^{TTC}* po hodowli na podłożu zawierającym 5-FOA, który wymusza pozbycie się plazmidu p416ADH-*TRZ1^{TTC}* z markerem pokarmowym URA, stają się *petite*.

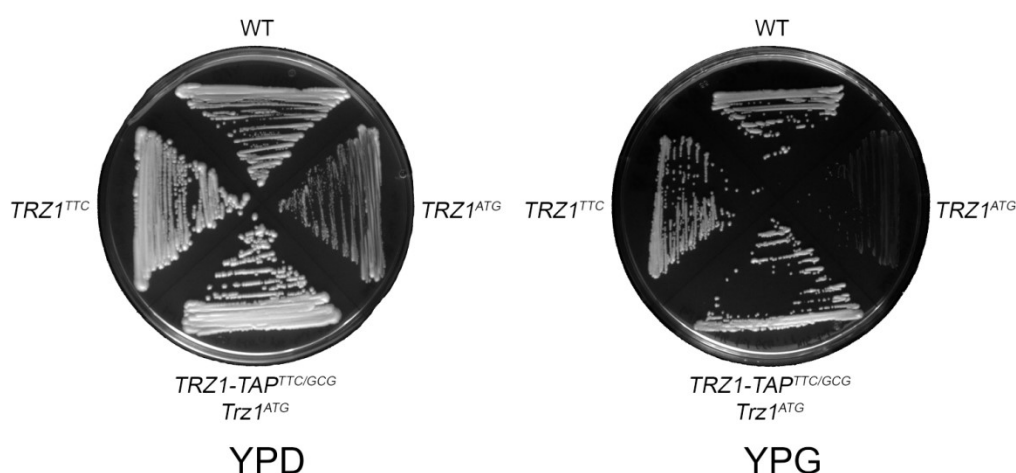


Fig 19. Testy wzrostowe różnych wariantów Trz1. Nieobecność izoformy wydłużonej powoduje fenotyp *petite* i niemożność wzrostu na podłożu, zawierającym glicerol (YPG). Testy wzrostowe na szalkach z glukozą (YPD) i glicerolem (YPG) kontroli (WT BY4741 oraz *TRZ1/trz1Δ* BY4743) oraz szczepów *trz1Δ* wyrażających warianty Trz1 opisane wcześniej.

Następnie potwierdzono lokalizację izoform Trz1. Podobnie jak przy weryfikacji lokalizacji poprzednich białek, przeprowadzono frakcjonowanie ekstraktów komórkowych szczepów drożdżowych, po czym wykryto białka znakowane znacznikiem TAP-tag techniką western blot. Zgodnie z poprzednimi obserwacjami, izoforma kanoniczna Trz1^{AUG}-TAP lokalizowała się jedynie w cytozolu, a produkty białkowe kodowane przez konstrukty *TRZ1^{TTC}-TAP* oraz *TRZ1^{TTC/GCG}-TAP* wykrywano w mitochondriach (Fig 20).

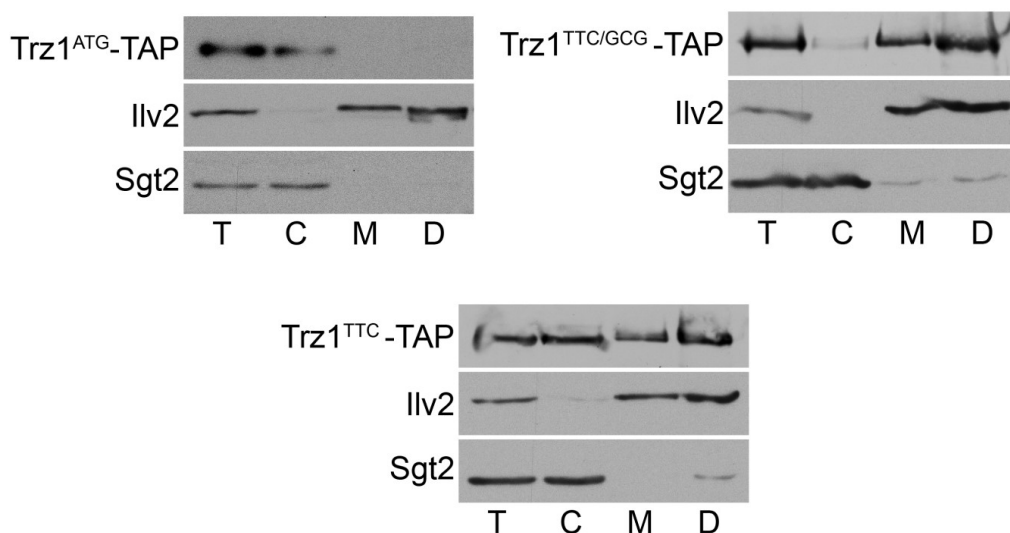


Fig 20. Frakcjonowanie szczepów, wyrażających różne warianty Trz1. Lokalizacja wariantów białka Trz1, w poszczególnych frakcjach. Do zlokalizowania wariantu Trz1^{ATG}-TAP użyto szczepu wyrażającego Trz1^{ATG}-TAP pod kontrolą promotora *GAL* z plazmidu z kolekcji MORF (Gelperin et al., 2005). Opis jak dla Fig 12.

Warto podkreślić, że izoformy Trz1^{UUC}-TAP i Trz1^{AUG}-TAP syntetyzowane z konstruktyw *TRZ1^{TTC}-TAP* znajdują się we wszystkich frakcjach, przy czym wariant mitochondrialny Trz1 migruje nieco wolniej, podobnie jak Trz1-TAP wyrażane z promotora natywnego w szczepie Trz1-TAP. Z kolei izoforma Trz1^{UUC} powstała w szczepie wyrażającym wariant *TRZ1^{TTC/GCG}-TAP* jest obecna przede wszystkim we frakcjach mitochondrialnych. Niewielka ilość cytozolowego Trz1 może odpowiadać Trz1^{UUC}, który nie został transportowany do mitochondriów, bądź izoformie Trz1 syntetyzowanej z kodonu GCG z niską wydajnością.

Obecność Trz1^{UUC} w cytozolu mimo wszystko jest zastanawiająca. W celu uzyskania dodatkowej kontroli, rozłożono tetrady *TRZ1/trz1Δ* transformowane jedynie plazmidem, zawierającym konstrukty *TRZ1^{TTC/GCG}-TAP*. Spodziewano się, że taki układ eksperymentalny będzie dla drożdży letalny, gdyż brak formy jądrowej Trz1 powoduje śmierć komórki. Okazało się jednak, że drożdże wyrażające jedynie formę wydłużoną Trz1^{UUG} zachowują normalny wzrost i fenotyp (wynik niepokazany). Aby upewnić się, że nie doszło do mutacji wstecznych, sprawdzono obecność mutacji ATG/GCG oraz delecji genu *TRZ1* w genomie metodą PCR na matrycy genomowego DNA szczepu *trz1Δ*+p416ADH-*TRZ1^{TTC/GCG}-TAP*, stosując specjalnie zaprojektowane startery. Wyniki potwierdziły obecność delecji genu *TRZ1* w genomie oraz nieobecność kodonu ATG w genie *TRZ1* kodowanym na plazmidzie (Fig 21).

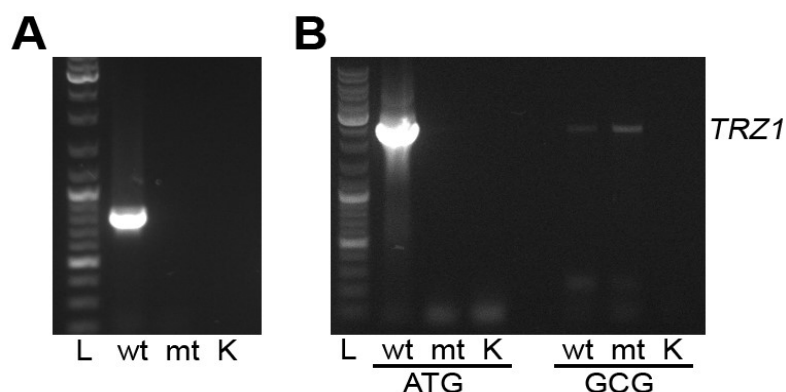


Fig 21. Potwierdzenie poprawności szczepu *trz1Δ*+p416ADH-*TRZ1*^{TTC/GCG}-TAP. (A) Potwierdzenie delecji genu *TRZ1* w genomie przy użyciu starterów F-Trz1-delta i R-Trz1-delta. (B) Potwierdzenie poprawności konstruktów plazmidowych. Elektroforeza w 1% żelu agarozowym w buforze 1x TAE z dodatkiem bromku etydydny. Startery lewe specyficzne do kodonu ATG genu *TRZ1* (F-Trz1-mutacja-ATG) oraz do kodonu GCG (F-Trz1-mutacja-GCG). Jako starter prawy wykorzystano R-Trz1. PCR przeprowadzono na matrycy genomowego DNA szczepu dzikiego (wt) oraz badanego *trz1Δ*+p416ADH-*TRZ1*^{TTC/GCG}-TAP (mt) wraz z kontrolą negatywną bez matrycy (K).

Następnie, aby sprawdzić, czy funkcje Trz1 w szczepie wyrażającym Trz1^{UUC} zostały zachowane, zbadano dojrzewanie prekursorów tRNA przy użyciu hybrydizacji typu northern blot. Do analizy wybrano dwa prekursory jądrowych tRNA (pre-tRNA_{IV}^{Lys1} i pre-tRNA_{VII}^{Phe}) o długich odcinkach prekursorowych po stronie 3', których biogeneza wymaga aktywności Trz1, a brak tej endonukleazy skutkuje widocznym defektem dojrzewania, oraz dwa dojrzałe mitochondrialne tRNA, których geny ulegają delecji z mitochondrialnego DNA przy braku mitochondrialnego wariantu Trz1 (Skowronek et al., 2014). Jako kontroli użyto RNA ze szczepu dzikiego BY4741 oraz *GAL::HA-TRZ1*, w którym ekspresja *TRZ1* była zatrzymana dzięki obecności glukozy (hodowla na pożywce YPD przez 12 godzin). W przypadku mitochondrialnych tRNA w badanym szczepie zaobserwowano fenotyp dziki, wskazujący na pełne zachowanie funkcji mitochondrialnych wariantu Trz1^{UUC}, co jest zgodne z fenotypem wzrostowym na podłożu z glicerolem (Fig 22A). Natomiast w przypadku prekursorów jądrowych występują defekty w dojrzewaniu tRNA, zbliżone do obserwowanych w szczepie *GAL::HA-TRZ1*, ale znacznie słabsze (Fig 22B). Uzyskany wynik oznacza, że sama forma wydłużona NTE-Trz1 jest w stanie zrekompensować brak wariantu podstawowego Trz1, ale tylko do pewnego stopnia. Jest możliwe, że w przypadku białka niezbędnego do przeżycia, pewna ilość wydłużonej formy Trz1 jest transportowana do jądra, gdzie pełni swoje funkcje. Nie można też wykluczyć, że kodon CGC, którym zastąpiono kanoniczny AUG, jest na niskim poziomie wykorzystywany jako kodon inicjujący. Wyjaśniałoby to obecność Trz1 w cytozolu w mutancie Trz1^{TTC/GCG}-TAP (Fig 20).

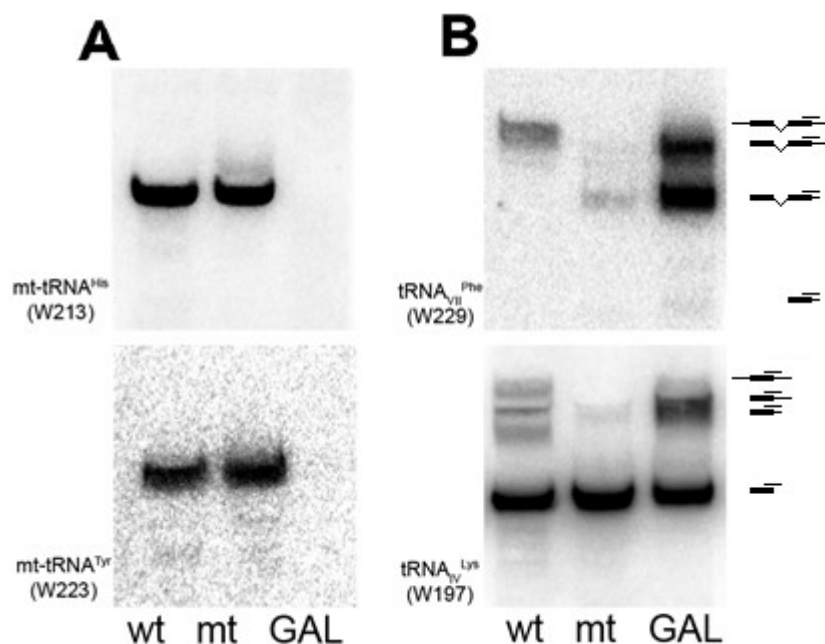


Fig 22. Dojrzewanie mitochondrialnych i jądrowych tRNA w szczepie *trz1Δ*+p416ADH-*TRZ1*^{TTC/GCG}-TAP. Analiza northern blot wybranych dojrzałych mitochondrialnych tRNA (A) i jądrowych pre-tRNA o długich sekwencjach prekursorowych na końcach 3' (B). Użyto frakcji małych RNA szczepu dzikiego (wt), *trz1Δ*+p416ADH-*TRZ1*^{TTC/GCG}-TAP (mt) oraz *GAL::HA-Trz1* poddanego 12-godzinnej deplecji na podłożu z glukozą (GAL). W nawiasach podano numery sond. Dla jądrowych dojrzałych i prekursorowych tRNA pokazano schemat tRNA i miejsce wiązania sondy.

W przypadku *Trz1* interesująca jest również wyraźna różnica w migracji między formą jądrową a mitochondrialną, wskazująca, że forma mitochondrialna może być modyfikowana lub mieć większą masę cząsteczkową od podstawowej, np. z powodu obecności NTE, który nie został odcięty przez MPP. Podobne zjawisko obserwuje się dla *Mtr4*, ale o ile w przypadku *Trz1* ewentualne różnice wielkości białek można wytłumaczyć stosunkowo długim wydłużeniem (54 aa), to dla *Mtr4* przewidziano bardzo krótkie NTE (3 i 6 aa). Dodatkowo, biorąc pod uwagę masę cząsteczkową *Mtr4* (122 kDa), wydaje się niemożliwe aby spowolnienie migracji było zależne jedynie od sekwencji aminokwasowej białka. Postanowiono więc sprawdzić, czy za obserwowany efekt odpowiedzialne są modyfikacje post-translacyjne *Mtr4*.

Z powodu bardzo krótkiej sekwencji NTE (IDYIRR) ilość miejsc potencjalnych modyfikacji jest bardzo ograniczona. Wybrano dwie możliwe modyfikacje post-translacyjne: glikozylację/glikację, prowadzącą do przyłączenia długich łańcuchów cukrowych (Gomes et al., 2015) oraz fosforylację, która jest jedną z najbardziej powszechnych modyfikacji post-translacyjnych białek (Humphrey et al., 2015). Frakcje cytozolową (C) oraz zagęszczoną mitochondrialną (D), uzyskane w wyniku frakcjonowania ekstraktu komórkowego szczepu *Mtr4*-TAP, poddano działaniu endoglikozydazy H (EndoH) odcinającej łańcuch cukrowy oraz

białkowej fosfatazy LPP faga Lambda, a następnie rozdzielono w 8% denaturującym żelu poliakrylamidowym SDS-PAGE w celu porównania różnic w migracji między formami cytozolową a mitochondrialną. Mtr4, znakowane na końcu C znacznikiem TAP-tag, wykrywano przy użyciu techniki western blot, korzystając z przeciwciał PAP, tak jak przy wcześniejszych analizach (Fig 23). Spodziewano się, że w przypadku zaistnienia danej modyfikacji, różnice w migracji między formą mitochondrialną a cytozolową nie będą widoczne po inkubacji z odpowiednim enzymem.

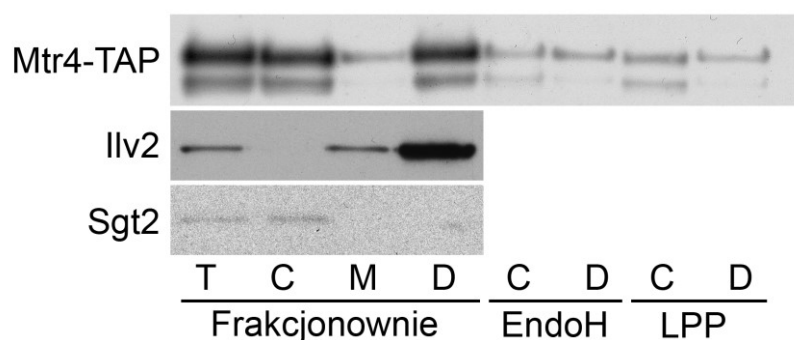


Fig 23. Analiza modyfikacji post-translacyjnych izoform Mtr4. Analiza western-blot. „Fracjonowanie” oznacza frakcje nietraktowane enzymami. EndoH – endoglikozydaza H, LPP – białkowa fosfataza faga Lambda. Do inkubacji z enzymami użyto frakcji cytozolowej (C) i zagęszczonej mitochondrialnej (D). Ilv2 i Sgt2 jak dla Fig 12.

Uzyskane wyniki nie pozwalają jednak jednoznacznie stwierdzić, którym z tych dwóch modyfikacji podlega mitochondrialny wariant Mtr4. Brak różnicy w migracji między formą cytozolową a mitochondrialną we frakcjach traktowanych EndoH wskazuje na glikozylację/glikację, jednakże może to być też artefakt, wynikający z zaburzeniami rozdziału elektroforetycznego np. z nierównomiernego rozkładu napięcia w żelu. Interesujące jest to, że sygnały po trawieniu EndoH zarówno dla formy cytozolowej jak i mitochondrialnej obserwuje się na tej samej wysokości jak dla frakcji mitochondrialnej, nietraktowanej żadnym z enzymów. Może to wskazywać na odwrotną sytuację niż zakładano na początku: modyfikacjom ulega kanoniczna forma białka, co powoduje zmianę właściwości fizyko-chemicznych i w efekcie zmianę prędkości migracji żelu (w tym przypadku przyspieszenie). Jednakże uzyskanie jednoznacznych wyników wymaga dalszych badań, łącznie z analizą spektrometrii mas oczyszczonej izoformy mitochondrialnej i cytozolowej. Uzyskane dane stanowią wyniki wstępne, będące początkiem badania nowych własności i funkcji białek NTE.

Podsumowując, białka MT-NTE nie tylko lokalizują się w mitochondriach, ale również pełnią w nich funkcje. Brak niektórych z nich (Trz1, Lsm1, Lsm6, Atg1 i Mtr4) wpływa znacząco na funkcjonowanie mitochondriów i wywołuje fenotyp defektu oddechowego.

5. Konserwacja NTE u wyższych eukariontów

Analizy bioinformatyczne jak i biochemiczne przeprowadzone do tej pory wskazały, jak powszechne jest zjawisko wykorzystywania niekanonicznych kodonów w inicjacji translacji wydłużonych izoform białek o lokalizacji mitochondrialnej u drożdży *S. cerevisiae*. Postanowiono sprawdzić, czy obecność NTE i MTS w ich obrębie jest konserwowana u wyższych eukariontów. Skupiono się na wybranych drożdżowych białkach z puli MTS⁺, które mają udokumentowane ludzkie homologi zdeponowane w bazie ENSEMBL (Cunningham et al., 2019). Sprawdzono dwa możliwe scenariusze: 1) udokumentowana lokalizacja mitochondrialna danego homologa w formie podstawowej. 2) lokalizacja mitochondrialna wydłużonej na końcu N izoformy, powstałej w wyniku startu syntezy białka z non-AUG aTIS. Potencjalne aTIS wykryto używając programu PreTIS (Reuter et al., 2016), a sygnał lokalizacji mitochondrialnej został wyliczony za pomocą MitoProt II, TargetP i MitoFates. Znalezione homologi dla 71 białek drożdżowych z puli MTS⁺, posiadających łącznie 115 ludzkich homologów (Tabela 13). Ich analiza ujawniła, że więcej niż połowa może lokalizować się w mitochondriach, z czego 26 białek drożdżowych posiada ludzkie homologi, których formy podstawowe mają lokalizację mitochondrialną, zaś 22 kierowane są do mitochondriów pod postacią NTE (Tabela 12).

Kategoria	Opis	Ilość	Lista białek drożdżowych
Bez MTS	Brak MTS w formie podstawowej i wydłużonej	24	Alg13, Avt5, Bur6, Cdc27, Crc1, Csl4, Fol2, Mrc1, Msn5, Nth2, Orc2, Per1, Pex13, Rbk1, Rnh1, Rpp0, Rps7a, Sea4, Sec18, Smd3, Syf1, Usb1, Vps15, Wrs1
Bez NTE	Brak przewidzianego NTE według PreTIS	16	Bur6, Cdc21, Cdc27, Csl4, Dip2, Erb1, Fol2, Mrc1, Nnr1, Nsg1, Nsg2, Nth2, Per1, Pex13, Rbk1, Syf1
MTS ⁺	Białka którym nadano parametr MTS ⁺ według takich samych reguł jak dla drożdży	18	Are1, Atg8, Car2, Cdc10, Chk1, Cit2, Emc3, Emc6, Hsl7, Npr2, Pby1, Pmt2, Prp43, Ptc1, Rpn6, Rps19a, Rrt12
MTS w formie podstawowej	MTS (MitoProt II, TargetP lub MitoFates) > 0,3	30	Acc1, Ala1, Bim1, Bpl1, Cdc14, Cdc21, Cdc31, Ceg1, Dip2, Erb1, Gdt1, Gpi8, Grs1, Has1, Lrp1, Mak5, Nnr1, Nsg1, Nsg2, Pcs60, Rad5, Rpl3, Rpl7b, Sif2, Sul1, Trz1, Ynr064c, Yor1, Ypk1, Ypk2
Posiadające MTS	MTS ⁺ + MTS w formie podstawowej	48	-

Tabela 12. Analiza konserwacji MTS na NTE wśród genów z puli MTS⁺ (białka drożdżowe). Otrzymane wyniki podzielono na kilka kategorii w zależności od obecności MTS. „Ilość” oznacza liczbę białek drożdżowych, które posiadają ludzkie homologi spełniające parametry, opisane dla danej kategorii. Wartość ta jest tożsama z ilością wypisanych białek w kolumnie „Lista białek drożdżowych”. We wierszu „Posiadające MTS” nie podano listy białek, by nie powtarzać danych.

Kategoria	Opis	Ilość	Lista białek ludzkich
Bez MTS	Brak MTS w formie podstawowej i wydłużonej	53	AARS, ALG13, CDC27, CLSPN, COPS2, DHX32, DRAP1, DUSP11, EXOSC1, GABARAPL1, GABARAPL2, GCH1, INSIG1, LGMN, MAP1LC3B, MAP1LC3B2, MIOS, NSF, ORC2, PEX13, PGAP3, PIK3R4, POMT1, PPM1E, PPM1F, PPM1L, RBKS, RNH1, RPLP0, RPS7, SBK1, SBK2, SBK3, SEPT1, SEPT2, SEPT4, SEPT9, SLC25A20, SLC25A29, SLC25A47, SLC38A7, SLC38A8, SNRPD3, SOAT1, SOAT2, TBL1XR1, TBL1Y, TREH, USB1, WARS, XAB2, XPO5, YJEFN3
Bez NTE	Brak przewidzianego NTE według PreTIS	43	AARS2, ACACB, BOP1, CDC14B, CDC27, CETN1, CETN2, CLSPN, COPS2, DRAP1, DUSP11, DUSP23, EXOSC1, GABARAPL2, GCH1, INSIG1, INSIG2, MAP1LC3B, MAP1LC3B2, MAP1LC3C, NAXE, PEX13, PGAP3, PIGK, PPM1E, PPM1F, RBKS, RPL3L, SBK2, SBK3, SEPT1, SGK3, SLC25A29, SLC25A47, SLC38A8, SOAT1, SOAT2, THOC3, TREH, TYMS, WDR3, XAB2, YJEFN3
MTS ⁺	Białka którym nadano parametr MTS ⁺ według takich samych reguł jak dla drożdży	17	CHEK1, CS, DQX1, EMC3, EMC6, GABARAP, ILKAP, NPRL2, OAT, PCSK9, POMT2, PRMT5, PSMD11, RPS19, SEPT3, SEPT7, TTL
MTS w formie podstawowej	MTS (MitoProt II, TargetP lub MitoFates) > 0,3	43	AARS2, AC000093.1, ACACA, ACSF2, ACSF3, BOP1, C1D, CDC14A, CDC14B, CETN1, CETN2, CETN3, CRTF, DDX18, DDX24, DGAT1, DUSP23, ELAC2, EPHX3, GARS, HLCS, HLTF, INSIG1, MAP1LC3A, MAP1LC3C, MAPRE1, MAPRE2, MAPRE3, NAXE, PIGK, RNGTT, RPL3, RPL3L, RPL7L1, SEPT12, SEPT5, SGK1, SGK2, SLC26A11, THOC3, TMEM165, TYMS, WDR3
Posiadające MTS	MTS ⁺ + MTS w formie podstawowej	60	-

Tabela 13. Analiza konserwacji MTS w obrębie NTE wśród genów z puli MTS⁺ (białka ludzkie). Otrzymane wyniki podzielono na kilka kategorii w zależności od obecności MTS. „Ilość” oznacza liczbę białek ludzkich, spełniających parametry, opisane dla danej kategorii. Wartość ta jest tożsama z ilością wypisanych białek w kolumnie „Lista białek ludzkich”. We wierszu „Posiadające MTS” nie podano listy białek, aby nie powtarzać danych

Wiele z otrzymanych w analizie konserwacyjnej białek jest kluczowe dla biogenezy mitochondriów, m.in. dojrzewania rybosomów czy syntezy mitochondrialnych kwasów tłuszczowych. Lokalizacja niektórych z nich, np. ACACA (homolog drożdżowego Acc1) czy ELAC2 (homolog Trz1), została potwierdzona eksperymentalnie (Brzezniak et al., 2011; Monteuis et al., 2017). 13 z nich (w tym 2 z grupy MTS⁺) znaleziono w bazie MitoCarta, gromadzącej informacje o białkach mitochondrialnych u ssaków (Calvo et al., 2016). Ponadto, wiadomo także, że ludzki eukariotyczny czynnik inicjujący eIF5A1, który nie został

ujęty w analizie, posiada izoformę, lokalizującą się w mitochondrium dzięki NTE (Pereira et al., 2016).

Wyniki te wskazują na to, że dane uzyskane dla drożdży mogą służyć jako punkt wyjściowy do identyfikacji ludzkich białek, których lokalizacja i funkcje mitochondrialne są dotychczas nieznane. Pozwoli to na lepsze zrozumienie działania mitochondriów oraz związanych z nimi chorób.

6. Podsumowanie

Dotychczas sądzono, że wykorzystanie kodonów non-AUG podczas inicjacji translacji w komórkach eukariotycznych do produkcji wydłużonych form białek jest zjawiskiem marginalnym. Przeprowadzone badania, zarówno bioinformatyczne jak i eksperymentalne, wskazują na to, że u drożdży *S. cerevisiae* translacja z niekanonicznych kodonów jest zdecydowanie bardziej powszechna niż się wydawało. Badania bioinformatyczne wykazały, że 2/3 białek w proteomie drożdżowym może zyskiwać wydłużenia na końcu N (NTE), a ponad 500 z nich w formie wydłużonej zyskuje lokalizację mitochondrialną. Potwierdzono eksperymentalnie lokalizację mitochondrialną dla 23 białek, a dla trzech z nich pokazano, że to forma wydłużona lokalizuje się w mitochondriach, podczas gdy kanoniczna znajduje się w cytozolu. Co istotne, izoformy białek powstałe w wyniku inicjacji translacji z kodonów non-AUG, poza lokalizacją mitochondrialną zyskują również funkcje w mitochondriach. Brak niektórych białek prowadzi do defektu oddechowego, tzw. fenotypu *petite*. Spośród 600 genów *pet*, ponad połowa to białka NTE, co wskazuje na funkcje mitochondrialne tych białek. Podsumowując, zjawisko wykorzystywania kodonów niekanonicznych w starcie syntezy białek może mieć znaczący wpływ na różnorodność proteomu mitochondrialnego.

Dyskusja

1. Synteza białek NTE zwiększa różnorodność proteomu

Wraz ze wzrostem złożoności organizmów zwiększa się też kompleksowość mechanizmów ekspresji genów. Dawny dogmat „jeden gen = jedno białko” jest od dłuższego czasu nieaktualny, szczególnie po poznaniu mechanizmów alternatywnej transkrypcji, splicingu czy translacji, które pozwalają na zwiększenie różnorodności proteomu przy zachowaniu tej samej ilości informacji w genomie (Mouilleron et al., 2016). W wyniku tych procesów powstaje wiele nowych izoform białek o odmiennej lokalizacji, stabilności czy funkcji (de Klerk i 'tHoen, 2015). Ponadto, stale odkrywane są nowe mechanizmy zwiększające różnorodność proteomu, jak chociażby synteza krótkich białek i peptydów na matrycy transkryptów wcześniej uznawanych za niekodujące (Pueyo et al., 2016; Plaza et al., 2017). Profilowanie rybosomów (Ribo-seq) umożliwia policzalne mapowanie miejsc translacji i pozwala na obserwację zdarzeń zachodzących poza głównym ORF, w tym szeroko pojęte użycie alternatywnych miejsc inicjacji translacji (aTIS) (Ingolia et al., 2009; Ingolia et al., 2011; Lee et al., 2012; Fritsch et al., 2012; Ingolia et al., 2014; Michel et al., 2014; Wang et al., 2016; Guenther et al., 2018). Zaskakujące jest, że kodony non-AUG stanowią połowę wszystkich aTIS i ponad 70% z tych położonych powyżej kanonicznego miejsca startu. Głównym mechanizmem wykorzystywania kodonów niekanonicznych w inicjacji translacji jest najprawdopodobniej przepuszczające skanowanie, dzięki któremu część rybosomów inicjuje syntezę białka z kodonów niepokrewnych w słabym kontekście Kozak, podczas gdy główny TIS znajduje się w kontekście optymalnym (Kozak, 2002). Wykorzystanie aTIS zależy również od długości 5'UTR, który może być niejednorodny w wyniku występowania licznych miejsc startu transkrypcji (Rojas-Duran i Gilbert, 2012; Kurihara et al., 2018).

Wysokoprzepustowe badania obejmujące wyliczenia bioinformatyczne, analizy profilowania rybosomów i badania proteomiczne, zwłaszcza mapowanie końców N białek, wskazują na powszechność niekanonicznych zjawisk translacyjnych w wyniku których powstają wydłużone bądź skrócone na końcu N izoformy białek o potencjalnie nowych właściwościach (Kochetov et al., 2005; Oyama et al., 2007; Wamboldt et al., 2009; Ivanov et al., 2011; Fritsch et al., 2012; Fournier et al., 2012; Kazak et al., 2013; Menschaert et al., 2013; Koch et al., 2014; Van Damme et al., 2014; Hsu et al., 2016; Wang et al., 2016; Sendoel et al., 2017; Zhang et al., 2017; Tang et al., 2017; Yeom et al., 2017; Na et al., 2018; Guenther et al., 2018; Spealman et al., 2018; van der Horst et al., 2019). Większość tych

analiz przeprowadzono dla organizmów wyższych, głównie komórek ludzkich (także nowotworowych), ale również dla myszy, drożdży, roślin (*Arabidopsis thaliana*), czy owadów (*Drosophila melanogaster*). Badania przedstawione w tej pracy wpisują się w najnowsze trendy badań dotyczących translacji i polegających na wyszukiwaniu alternatywnych ścieżek syntezy białek w skali globalnej.

Analizy bioinformatyczne dostępnych danych z profilowania rybosomów dla *S. cerevisiae* potwierdzają powszechność wykorzystywania kodonów non-AUG w inicjacji translacji (Monteiuss, Miścicka et al., 2019). Okazuje się, że 10% przewidzianych aTIS jest wykorzystywana w normalnych warunkach, a można założyć, że w warunkach stresowych ich liczba znacznie wzrośnie. Analizy te wskazują, że wydłużone izoformy powstające z dodatkowego aTIS położonego powyżej standardowego kodonu AUG są wyrażane dla 123 białek MTS⁺ i 34 białek MTS⁻, dotyczy to także wielu z przebadanych eksperymentalnie białek, np. Trz1, Gpp1, Gpp2, Trm3 czy Lsm4. Sprawdzono również występowanie wewnętrznych kodonów AUG położonych poniżej kodonu kanonicznego, których wykorzystanie może prowadzić do powstania skróconych form białek (ang. *N-terminally truncated*, NTT). Wyliczenia bioinformatyczne wykazały 4 049 takich przypadków w całym genomie drożdżowym, a dane Ribo-seq potwierdziły syntezę 301 białek NTT. Przeprowadzono również analizę kontestów Kozak (sekwencja od -6 nukleotydu do +4) dla każdej przewidzianej izoformy. Wbrew oczekiwaniom okazało się, że sekwencja Kozak nie jest jednak szczególnie istotna dla wyboru kodonów inicjujących u drożdży *S. cerevisiae* (Monteiuss, Miścicka et al., 2019).

2. Czynniki wpływające na wybór aTIS u drożdży

Wykorzystywanie aTIS jest zdecydowanie lepiej zbadane w komórkach ludzkich niż u drożdży, stąd o wiele więcej wiadomo o ludzkich czynnikach regulujących to zjawisko (Sriram et al., 2018). W przypadku drożdży, ostatnie badania pokazały, że wybór kodonów spokrewnionych na miejsca inicjacji translacji zarówno w przypadku ekspresji uORF jak i NTE jest kontrolowany przez helikazę RNA Ded1 z rodziny DEAD-box (Guenther et al., 2018). Warto wspomnieć, że wybór aTIS może być szczególnie istotny podczas stresu (np. oksydacyjnego, odpowiedzi na niesfałdowane białka czy podczas głodzenia) lub w trakcie cyklu komórkowego (np. mejozy), kiedy dochodzi do przebudowania proteomu w odpowiedzi na zmianę warunków komórkowych (Ingolia et al., 2009; Gerashchenko et al., 2012; Andreev et al., 2015; Pakos-Zebrucka et al., 2016; Starck et al., 2016; Van Dalfsen

et al., 2018; Guenther et al., 2018; Crawford i Pavitt, 2019). Relatywnie dobrze poznana jest rola uORF w różnorodnych aspektach regulacji translacji, jednak znaczenie form białkowych wydłużonych na końcu N wciąż nie jest w pełni poznane.

Najnowsze badania nad strukturą rybosomów wskazują na ich heterogenność pod kątem kompozycji białkowej, co znacząco wpływa na ich funkcje. Coraz więcej mówi się o wyspecjalizowanych rybosomach (Emmott et al., 2019), które powstają *de novo* do pełnienia określonych funkcji (np. syntezy białek mitochondrialnych), albo są generowane poprzez przebudowanie istniejących rybosomów w warunkach stresowych (Ferretti et al., 2017; Samir et al., 2018; Segev i Gerst, 2018; Ferretti et al., 2018). Białko Rps26, będące składnikiem małej podjednostki rybosomu, bierze udział w rozpoznaniu przez kompleks preinicjujący (PIC) kontekstu Kozak w cząsteczkach mRNA, które kodują kluczowe geny szlaków metabolicznych. Rybosomy pozbawione Rps26, powstające w warunkach stresowych, są mniej wrażliwe na zmiany w sekwencji Kozak i mogą inicjować translację z kodonów położonych w znacznie słabszych kontekstach, regulując w ten sposób odpowiedź na stres solny czy zasadowy odczyn pH (Ferretti et al., 2017; Ferretti et al., 2018).

Budowa PIC ma znaczenie również przy wyborze kodonów niekanonicznych. Kompleks preinicjujący ulega przebudowaniu po zakończeniu skanowania (z pozycji P_{OUT}) i wyborze kodonu inicjującego zarówno kanonicznego, jak i non-AUG (na pozycję P_{IN}) (Bowen et al., 2015; Llácer et al., 2018). Niezwykle istotny jest udział poszczególnych czynników inicjujących na danym etapie inicjacji translacji. Kluczowym czynnikiem, wpływającym na wybór kodonów niepokrewnych jest eIF1. Pojedyncze mutacje w pętli 1 i 2 eIF1 znacząco zwiększają częstość wyboru na miejsce startu kodonu UUG i stabilizację całego kompleksu nawet w słabym kontekście Kozak (Thakur i Hinnenusch, 2018).

3. W wyniku inicjacji translacji z kodonów startu non-AUG powstaje wiele białek o lokalizacji mitochondrialnej

Na podstawie przeprowadzonych badań można przedstawić następujący model opisujący lokalizację białek MT-NTE (mitochondrialnych NTE). W przypadkach, gdy w obrębie mRNA znajdują się dwa kodony startu translacji – kanoniczny AUG (TIS) i położony powyżej w tej samej ramce odczytu niekanoniczny non-AUG (aTIS) – syntetyzowane są dwie izoformy białkowe, kanoniczna i wydłużona, zawierająca wydłużenie na końcu N (ang. *N-terminal extension*, NTE). Alternatywne miejsce startu translacji jest używane w znacznie mniejszym stopniu niż TIS. NTE może zawierać nowy sygnał lokalizacji

komórkowej, którym w przypadku białek MT-NTE jest MTS, dzięki któremu wydłużona izoforma lokalizuje się w innym organelum niż forma kanoniczna (Fig 25).

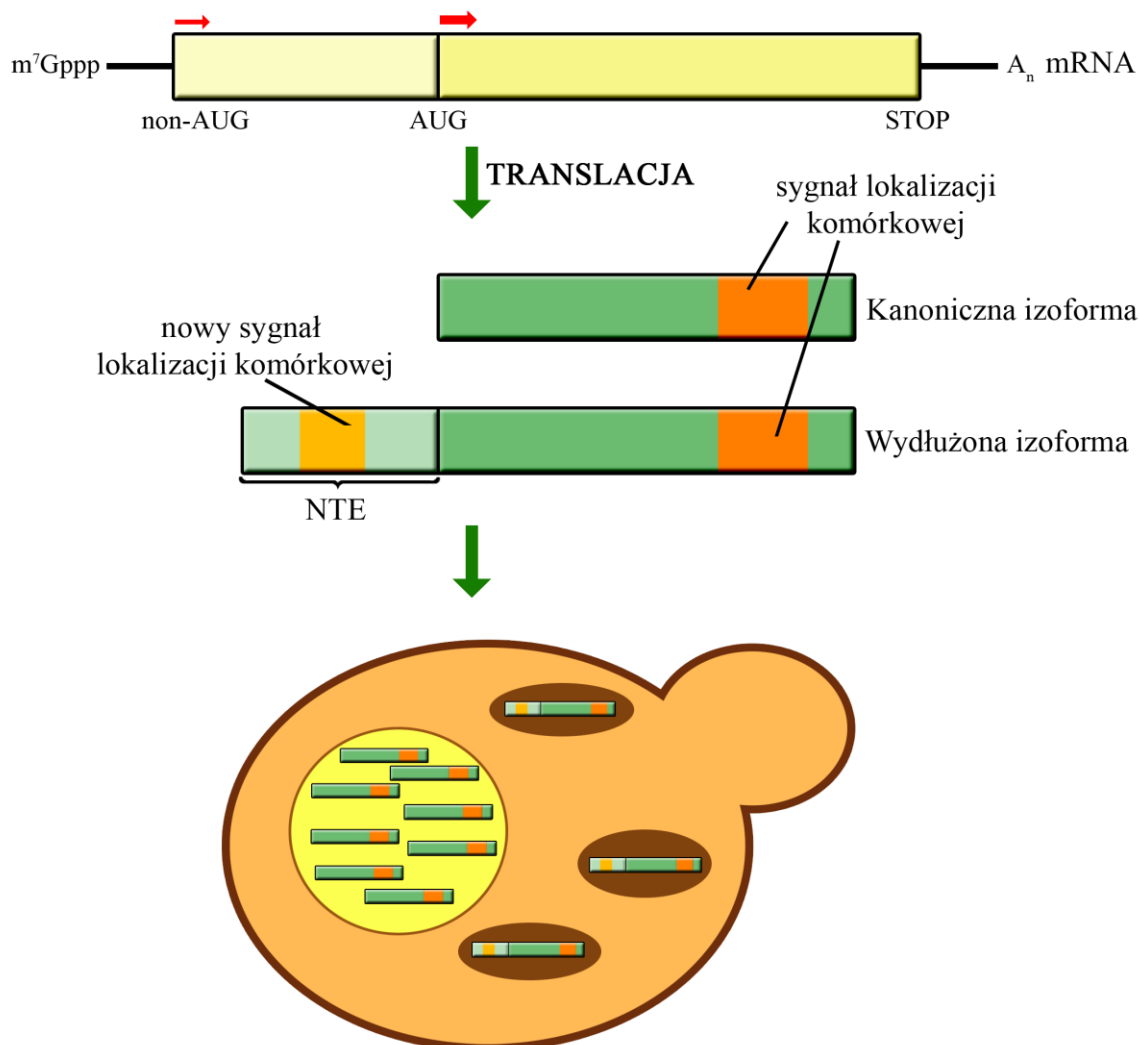


Fig 25. Schemat podwójnej lokalizacji białek u drożdży wynikającej z użycia dwóch TIS. W wyniku startu translacji z niekanonicznego i kanonicznego kodonu startu powstają dwie izoformy białka o różnej lokalizacji, wydłużona i kanoniczna. Na schemacie komórki drożdżowej pokazano przykładową lokalizację izoformy kanonicznej w jądrze komórkowym (oznaczonym na żółto), a wydłużonych izoform w mitochondriach (brązowe elipsy). Legenda: żółty prostokąt - ORF w obrębie mRNA; zielone prostokąty - białka powstałe po translacji; jaśniejsze kolory - wydłużenie na końcu N (NTE, ang. *N-terminal extension*) powstałe w wyniku niekanonicznej translacji; pomarańczowe prostokąty - sygnały lokalizacji komórkowej.

Ogólnie białka NTE mogą zyskiwać wiele różnych funkcji, lecz do tej pory przewidziano głównie lokalizację mitochondrialną, którą w kilku przypadkach potwierdzono eksperymentalnie u ssaków, roślin i drożdży (Natsoulis et al., 1986; Chaton et al., 1988; Tang et al., 2004; Chang i Wang, 2004; Kochetov et al., 2005; Wamboldt et al., 2009; Kazak et al., 2013; Suomi et al., 2014; Kritsiligkou et al., 2017). Przeprowadzone analizy *in silico*, poparte danymi z profilowania rybosomów, wskazują na to, że w komórkach drożdżowych tysiące

genów posiada więcej niż jeden TIS, których wykorzystanie prowadzi do powstania setek izoform białkowych, w tym wielu mitochondrialnych. Potwierdzono eksperymentalnie lokalizację mitochondrialną większości wybranych białek, wskazując na istotny udział NTE i znajdującego się w jego obrębie MTS w tym procesie (Monteuuis, Miścicka et al., 2019). Warto dodać, że wydłużenia na końcu N mogą nieść również inne sygnały lokalizacji komórkowej jak i motywy strukturalne czy funkcjonalne domeny. Tak jest w przypadku długiej izoformy Kgd4, której NTE powstałe w wyniku inicjacji z kodonu UUG bierze udział w formowaniu domeny kotwiczącej do błony mitochondrialnej (Heublein, 2014; Heublein et al., 2014; Heublein et al., 2019).

Wyniki przedstawione w pracy wskazują również białka, które transportowane są do mitochondriów pomimo braku wykrywalnego MTS zarówno w formie podstawowej jak i wydłużonej (Air2, Naf1, Nrd1, Rna14, Rna15). Pozwala to wysnuć hipotezę o możliwości transportu niektórych białek do mitochondriów pod postacią całych kompleksów. Białko Air2 jest składnikiem kompleksu TRAMP (ang. *Trf4/Air2/Mtr4p Polyadenylation complex*), zaangażowanego w obróbkę i degradację RNA w jądrze komórkowym (Vanáčová et al., 2005; LaCava et al., 2005; Jia et al., 2011). Kompleks TRAMP składa się z białek Trf4/Trf5, Air1/Air2 oraz Mtr4, dla którego potwierdzono w tej pracy lokalizację mitochondrialną za pośrednictwem MTS znajdującego się w obrębie NTE. Ponadto, mimo że Trf4 nie ma przewidzianej formy NTE, według MitoProt II posiada bardzo silny sygnał lokalizacji mitochondrialnej (0,914; 0,201 według TargetP i 0,000 według MitoFates). Możliwe, że pozostałe składniki kompleksu TRAMP, włączając zbadane Air2 i Trf4, kierowane są do mitochondriów jako cały kompleks za pośrednictwem Mtr4. Kolejne białko Naf1, niezbędne do dojrzewania niekodujących RNA takich jak rRNA, tRNA i snoRNA (Fatica et al., 2002), oddziałuje fizycznie z białkiem Cbf5 (Fatica et al., 2002; Yang et al., 2005), które należy do puli MTS⁺ i jest zakwalifikowane do klasy 4 według badań proteomu mitochondrialnego (Morgenstern et al., 2017). Z kolei Nrd1 jest składnikiem kompleksu NNS (Nrd1, Nab3, Sen1), który wraz z egzosomem pośredniczy w formowaniu końca 3' niektórych niekodujących RNA (Vasiljeva et al., 2008). Nrd1 i Nab3 zostały zidentyfikowane w proteomie mitochondrialnym i zaliczone, odpowiednio, do klasy 3 i 4 (Morgenstern et al., 2017). Co istotne, kompleks NNS oddziałuje z kompleksem TRAMP (Vasiljeva et al., 2008) i również może być kierowany do mitochondriów za pośrednictwem Mtr4. W przypadku Rna14 i Rna15, które są składnikami czynnika cięcia i poliadenylacji CF I odpowiedzialnego za formowanie końca 3' cząsteczek mRNA (Gross i Moore, 2001),

wiadomo jest, że Rna14 jest obecne w mitochondriach, ale mechanizm lokalizacji nie jest znany (Rouillard et al., 2000; Brendolise et al., 2002). Samo białko Rna15 ma dość wysoki MTS w formie podstawowej (0,577 według TargetP).

Możliwe jest, że w przypadku innych kompleksów również zachodzą podobne zależności. Ciekawym przypadkiem są białka Lsm, które tworzą dwa kompleksy i biorą udział w splicingu pre-mRNA w jądrze komórkowym (kompleks złożony z białek Lsm2-8) lub w degradacji mRNA w cytoplazmie (kompleks złożony z białek Lsm1-7) (Tharun, 2009). Interesujące jest, że wszystkie białka z rodziny Lsm mają przewidziany NTE, a trzy z nich należą do kategorii MTS⁺ (Tabela 14). Dla Lsm1 oraz Lsm4 potwierdzono lokalizację mitochondrialną w tej pracy (Fig 12-13), a dla Lsm1 dodatkowo pokazano, że tylko forma wydłużona lokalizuje się w mitochondriach (Fig 14). W przypadku Lsm6 zaobserwowano niewielki defekt oddechowy w szczepie *lsm6Δ*, a *LSM1* jest znanym genem *pet* (Dimmer et al., 2002; Merz i Westermann, 2009). Kolejne przeprowadzone przez nas badania potwierdziły lokalizację mitochondrialną dla Lsm1, Lsm3, Lsm4 oraz Lsm7, i wykazały, że Lsm8 nie jest obecne w tym przedziale komórkowym, natomiast ustalenie lokalizacji Lsm2, Lsm5 i Lsm6 wymaga dalszych badań (Edwalds-Gilbert, Miścicka, Kufel, dane nieopublikowane). Możliwe jest więc, że w mitochondrium istnieje alternatywny kompleks Lsm, lecz jego kompozycja i rola nie zostały jeszcze zbadane.

Białko	Kanoniczny MTS (średnia MP/TP/MF)	NTE MTS (średnia MP/TP/MF)	Długość NTE	Klasa (Morgenstern et al., 2017)
Lsm1	0,082	0,291	5, 21	Klasa 3
Lsm2	0,076	0,043	10, 19, 23	Klasa 3
Lsm3	0,036	0,151	14, 16, 22	Klasa 3
Lsm4	0,042	0,190	6, 11, 17, 18	-
Lsm5	0,111	0,076	11, 13	Klasa 3
Lsm6	0,078	0,399	1, 13, 18, 20, 37, 47	Klasa 3
Lsm7	0,036	0,100	5, 11	-
Lsm8	0,080	0,097	19, 26	-

Tabela 14. Składniki kompleksu Lsm wraz z NTE i wyliczonym MTS. Opis analogicznie do Tabeli 9. Na żółto zaznaczono białka zaliczone do kategorii MTS⁺.

4. Funkcje mitochondrialnych izoform białkowych

Wiele białek MTS⁺, o potencjalnej lokalizacji mitochondrialnej, odgrywa istotną rolę w obróbce RNA (Tabela 15). Według SGD ponad 20 białek z puli MTS⁺ jest zaangażowanych w różne etapy dojrzewania rybosomów. Podczas gdy dotychczas opisano ponad 200 czynników niezbędnych do formowania rybosomów cytoplazmatycznych,

biogeneza rybosomów mitochondrialnych wciąż jest słabo poznana i tylko udział nielicznych białek jest dobrze opisany (Fernández-Pevida et al., 2015).

Kolejnymi ciekawymi białkami MTS⁺ związanymi z metabolizmem RNA są te zaangażowane w transkrypcję przez polimerazę II RNA, splicing pre-mRNA oraz obróbkę i degradację RNA. W tym momencie trudno jest przewidzieć dokładną funkcję tych białek w metabolizmie mitochondrialnych RNA, ale przynajmniej kilka z nich ma udokumentowaną aktywność mitochondrialną. Wśród nich są aminoacylasy tRNA: Hts1, Vas1, Ala1 i Grs1, pełniące analogiczne funkcje w cytoplazmie (Tang et al., 2004; Chang i Wang, 2004; Natsoulis et al., 1986; Chatton et al., 1988), oraz Trz1 i Sen2, które biorą udział w dojrzewaniu mitochondrialnych transkryptów (Yoshihisa et al., 2003; Skowronek et al., 2014). Rybonukleaza Rnh1 (RNaza H1), rozpoznająca i tnąca hybrydy RNA:DNA, jest zaangażowana w usuwanie pętli R w mitochondriach (El Hage et al., 2014), a spliceosomalna helikaza RNA Prp43 z rodziny DEAH-box bierze udział w procesie apoptozy w mitochondriach (Heininger, 2016). Z kolei Jsn1, cytoplazmatyczne białko wiążące RNA z rodziny Puf, pośredniczy w lokalizowaniu kompleksu Arp2/3 w mitochondriach (Fehrenbacher et al., 2005). Warto podkreślić, że ludzkie homologe tych czynników, a dokładnie endonukleaza tRNA RNaza Z_L, RNaza H1 i helikaza DHX15, mają podwójną jądrowo-mitochondrialną lokalizację i znane funkcje mitochondrialne (Cerritelli et al., 2003; Rossmanith, 2011; Mosallanejad et al., 2014; Holmes et al., 2015). Pełna lista białek MTS⁺ zaangażowanych w metabolizm RNA została przedstawiona w Tabeli 15.

Funkcja	Geny
Transkrypcja	<i>MED4, MED8, PGD1, ROX3, SOH1, TAH1</i>
Splicing	<i>CWC23, DRN1, NPL3, NTC20, PRP22, PRP39, PRP43, PUS7, SYF1, SYF2, USB1</i>
Biogeneza i funkcje rybosomów	<i>CBF5, DBP9, DIB1, DIP2, ERB1, ESF1, HAS1, MAK5, NOP19, NOP9, PNO1, PUF4, RDS3, RPL3, RPL7B, RRP36, RRP8, SPB4, TRS4, UTP10, UTP4</i>
Dojrzewanie, splicing i modyfikacje tRNA	<i>ALA1, ATSI, DUS1, GRS1, KAE1, MES1, MSN5, PCC1, PPM2, PTC1, SEN15, SEN2, SOL1, STP3, TRM3, TRM82, TRZ1, WRS1, YDR341C</i>
Translacja	<i>DED1, DOM34, ELP6, SFL1, SLH1, TIF1</i>
Degradacja RNA	<i>CEG1, FIR1, LSM1, LSM4, LSM6, MMS1, MOT2, MTR4, NDM4, POP2, STO1, SUP35</i>
Egzosom	<i>CSL4, LRP1, RRP45, SKI3</i>
Inne	<i>BFR1, CTR9, EAR1, HPC2, JSN1, MOT1, NAB6, NET1, PBY1, RIE1, RNH1, RPB8, RPO31, RTC3, SHE2, SUB1</i>

Tabela 15. Geny, kodujące białka MTS⁺ zaangażowane w metabolizm RNA. W kolumnie „Geny” wypisano te, zaliczone do kategorii MTS⁺, pogrupowane według określonej funkcji w obróbce RNA (kolumna „Funkcja”).

Nie tylko białka zaangażowane w metabolizm RNA lokalizują się i pełnią funkcje w mitochondriach. Również enzymy szlaku syntezy kwasów tłuszczowych, takie jak Ayr1

(bifunkcyjna lipaza triacyloglycerylowa i reduktaza 1-acyl-DHAP), Hfa1 (karboksylaza acetyl-CoA), Crc1 (transporter karnityny) czy Hst4 (NAD(+)-zależna deacetylaza białkowa) mają udokumentowaną zarówno lokalizację jak i funkcję mitochondrialną (van Roermund et al., 1999; Zahedi et al., 2006; Soumi et al., 2014; Madsen et al., 2015). W mitochondriach gromadzą się także liczne białka odpowiedzi na stres (Hyr1, Trr1, Uth1, Hst4) (Pedrajas et al., 1999; Kissová et al., 2004; Trotter i Grant, 2005; Madsen et al., 2015; Kritsiligkou et al., 2017), podobnie jak transportery i białka pośredniczące w transporcie międzyorganellarnym (Pfk1, Yia6, Ymc1, Num1) (Graf et al., 1993; Brandina et al., 2006; Todisco et al., 2006; Lackner et al., 2013), czy białka zaangażowane w autofagię (Atg27) (Yen et al., 2007).

Wstępne analizy ludzkich homologów drożdżowych białek MTS⁺ sugerują, że wykorzystanie aTIS do generowania izoform białkowych o odmiennej lokalizacji nie jest charakterystyczne jedynie dla drożdży, tym bardziej, że w komórkach ludzkich istnieje znacznie więcej możliwości syntezy alternatywnych izoform białkowych. W związku z tym, zastawana metoda badawcza umożliwia częściowe określenie potencjalnej lokalizacji mitochondrialnej białek, które nie są anotowane jako mitochondrialne. Można więc to podejście wykorzystać do identyfikacji czynników odpowiadających za choroby mitochondrialne, zwłaszcza że izoformy NTE-MTS są z reguły wyrażane na niskim poziomie, przez co są trudniejsze do zidentyfikowania jako patogenne.

Podsumowując, uzyskano dane potwierdzające wielkość i złożoność proteomu mitochondrialnego u drożdży *S. cerevisiae*. Aby móc w pełni poznać to kompleksowe zjawisko konieczna jest dalsza analiza białek niemitochondrialnych o potencjalnej lokalizacji i funkcji w mitochondriach, a następnie zbadanie ich aktywności w mitochondriach. Warto pamiętać, że wydłużenia na końcu N przewidziane są dla ponad 4 000 białek drożdżowych, nie tylko potencjalnie mitochondrialnych. Bardziej odkrywcze może być zbadanie pozostałych właściwości nadanych przez NTE, takich jak sygnały kierujące do innych przedziałów komórkowych, dodatkowe domeny funkcjonalne i strukturalne, miejsca modyfikacji post-translacyjnych czy elementy decydujące o stabilności białek. Szczególnie istotną kwestią, obecnie bardzo słabo rozpoznaną, jest całokształtowy wpływ tego "ukrytego" do tej pory proteomu na efektywne funkcjonowanie komórek eukariotycznych.

Literatura

- Acland, P., Dixon, M., Peters, G. and Dickson, C. (1990) Subcellular fate of the int-2 oncoprotein is determined by choice of initiation codon. *Nature* 343, 662–665.
- Altmann, K. and Westermann, B. (2005) Role of essential genes in mitochondrial morphogenesis in *Saccharomyces cerevisiae*. *Mol Biol Cell*. 16(11):5410-7.
- Altmann, K., Dürr, M. and Westermann, B. (2007) *Saccharomyces cerevisiae* as a model organism to study mitochondrial biology: general considerations and basic procedures. *Methods Mol Biol*. 372:81-90.
- Anaganti, S., Hansen, J.K., Ha, D., Hahn, Y., Chertov, O., Pastan, I. and Bera, T.K. (2009) Non-AUG translational initiation of a short CAPC transcript generating protein isoform. *Biochem Biophys Res Commun*. 380(3):508-13.
- Andreev, D.E., O'Connor, P.B., Fahey, C., Kenny, E.M., Terenin, I.M., Dmitriev, S.E., Cormican, P., Morris, D.W., Shatsky, I.N. and Baranov, P.V. (2015) Translation of 5' leaders is pervasive in genes resistant to eIF2 repression. *Elife*, 4, e03971.
- Andreev, D.E., O'Connor, P.B., Loughran, G., Dmitriev, S.E., Baranov, P.V. and Shatsky, I.N. (2017) Insights into the mechanisms of eukaryotic translation gained with ribosome profiling. *Nucleic Acids Res*. 45(2):513-526.
- Arnaud, E., Touriol, C., Boutonnet, C., Gensac, M.C., Vagner, S., Prats, H. and Prats, A.C. (1999) A new 34-kilodalton isoform of human fibroblast growth factor 2 is cap dependently synthesized by using a non-AUG start codon and behaves as a survival factor. *Mol Cell Biol*. 19(1):505-14.
- Backes, S., Hess, S., Boos, F., Woellhaf, M.W., Gödel, S., Jung, M., Mühlhaus, T. and Herrmann, J.M. (2018) Tom70 enhances mitochondrial preprotein import efficiency by binding to internal targeting sequences. *J Cell Biol*. 217(4):1369-1382.
- Baranov, P.V., Fayet, O., Hendrix, R.W. and Atkins, J.F. (2006) Recoding in bacteriophages and bacterial IS elements. *Trends Genet*. 22(3):174-81.
- Bekaert, M., Firth, A.E., Zhang, Y., Gladyshev, V.N., Atkins, J.F. and Baranov P.V. (2010) Recode-2: new design, new search tools, and many more genes. *Nucleic Acids Res*. 38(Database issue):D69-74.
- Belcourt, M.F. and Farabaugh, P.J. (1990) Ribosomal frameshifting in the yeast retrotransposon Ty: tRNAs induce slippage on a 7 nucleotide minimal site. *Cell*. 1990 62(2):339-52.
- Bertram, G., Innes, S., Minella, O., Richardson, J. and Stansfield, I. (2001) Endless possibilities: translation termination and stop codon recognition. *Microbiology*. 147(Pt 2):255-69.
- Bowen, A.M., Musalgaonkar, S., Moomau, C.A., Gulay, S.P., Mirvis, M. and Dinman, J.D. (2015) Ribosomal protein uS19 mutants reveal its role in coordinating ribosome structure and function. *Translation (Austin)*. 3(2):e1117703.
- Brameier, M., Krings, A. and MacCallum, R.M. (2007) NucPred--predicting nuclear localization of proteins. *Bioinformatics*. 23(9):1159-60.
- Brandina, I., Graham, J., Lemaitre-Guillier, C., Entelis, N., Krasheninnikov, I., Sweetlove, L., Tarassov, I. and Martin, R.P. (2006) Enolase takes part in a macromolecular complex associated to mitochondria in yeast. *Biochim Biophys Acta*. 1757(9-10):1217-28.
- Brar, G.A., Yassour, M., Friedman, N., Regev, A., Ingolia, N.T. and Weissman, J.S. (2012) High-resolution view of the yeast meiotic program revealed by ribosome profiling. *Science* 335(6068):552-7.
- Brendolise, C., Rouillard, J.M., Dufour, M.E., Lacroute, F. (2002) Expression analysis of RNA14, a gene involved in mRNA 3' end maturation in yeast: characterization of the rna14-5 mutant strain. *Mol Genet Genomics*. 267(4):515-25.
- Brown, A. and Shao, S. (2018) Ribosomes and cryo-EM: a duet. *Curr Opin Struct Biol*. 52:1-7.
- Brown, J. D. and Ryan, M. D. (2010) Ribosome “skipping”: “stop-carry on” or “StopGo” translation. In *Recoding: Expansion of Decoding Rules Enriches Gene Expression*, pp. 101–121. Edited by J. F. Atkins & R. F. Gesteland. Heidelberg: Springer.
- Brzezniak, L.K., Bijata, M., Szczesny, R.J. and Stepień, P.P. (2011) Involvement of human ELAC2 gene product in 3' end processing of mitochondrial tRNAs. *RNA Biol*. 8(4):616-26.
- Calvo, S.E., Clauser, K.R. and Mootha, V.K. (2016) MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. *Nucleic Acids Res*. 44(D1):D1251-7.

- Camasses, A., Bragado-Nilsson, E., Martin, R., Séraphin, B. and Bordonné, R. (1998) Interactions within the yeast Sm core complex: from proteins to amino acids. *Mol Cell Biol.* 18(4):1956-66.
- Castilho-Valavicius, B., Yoon, H. and Donahue, T.F. (1990) Genetic characterization of the *Saccharomyces cerevisiae* translational initiation suppressors *sui1*, *sui2* and *SUI3* and their effects on *HIS4* expression. *Genetics* 124: 483–495.
- Cerritelli, S.M., Frolova, E.G., Feng, C., Grinberg, A., Love, P.E. and Crouch, R.J. (2003) Failure to produce mitochondrial DNA results in embryonic lethality in *Rnaseh1* null mice. *Mol Cell*, 11, 807–815.
- Chang, A.Y., Chau, V.W.Y., Landas, J.A. and Pang, Y. (2017) Preparation of calcium competent *Escherichia coli* and heat-shock transformation. *JEMI Methods* 1:22-27.
- Chang, C.P., Chen, S.J., Lin, C.H., Wang, T.L. and Wang, C.C. (2010) A single sequence context cannot satisfy all non-AUG initiator codons in yeast. *BMC Microbiol.*, 10, 188.
- Chang, K.J. and Wang, C.C. (2004) Translation initiation from a naturally occurring non-AUG codon in *Saccharomyces cerevisiae*. *J. Biol. Chem.*, 279, 13778-13785.
- Chatton, B., Walter, P., Ebel, J.P., Lacroute, F. and Fasiolo, F. (1988) The yeast *VAS1* gene encodes both mitochondrial and cytoplasmic valyl-tRNA synthetases. *J. Biol. Chem.*, 263, 52-57.
- Chen, J., Coakley, A., O'Connor, M., Petrov, A., O'Leary, S.E., Atkins, J.F. and Puglisi, J.D. (2015) Coupling of mRNA Structure Rearrangement to Ribosome Movement during Bypassing of Non-coding Regions. *Cell*. 163(5):1267-1280.
- Chen, S.J., Lin, G., Chang, K.J., Yeh, L.S. and Wang, C.C. (2008) Translational efficiency of a non-AUG initiation codon is significantly affected by its sequence context in yeast. *J Biol Chem*. 283(6):3173-80.
- Chen, X.J. and Clark-Walker, G.D. (2000) The *petite* mutation in yeasts: 50 years on. *Int Rev Cytol.* 194:197-238.
- Chen, Y., Beck, A., Davenport, C., Shattuck, D. and Tavtigian, S.V. (2005) Characterization of TRZ1, a yeast homolog of the human candidate prostate cancer susceptibility gene ELAC2 encoding tRNase Z. *BMC Mol. Biol.*, 6, 12.
- Cheng, Z., Otto, G.M., Powers, E.N., Keskin, A., Mertins, P., Carr, S.A., Jovanovic, M. and Brar, G.A. (2018) Pervasive, Coordinated Protein-Level Changes Driven by Transcript Isoform Switching during Meiosis. *Cell*. 172(5):910-923.e16.
- Cherry, J.M., Hong, E.L., Amundsen, C., Balakrishnan, R., Binkley, G., Chan, E.T., Christie, K.R., Costanzo, M.C., Dwight, S.S., Engel, S.R., Fisk, D.G., Hirschman, J.E., Hitz, B.C., Karra, K., Krieger, C.J., Miyasato, S.R., Nash, R.S., Park, J., Skrzypek, M.S., Simison, M., Weng, S. and Wong, E.D. (2012) *Saccharomyces* Genome Database: the genomics resource of budding yeast. *Nucleic Acids Res.* 40 (Database issue):D700-5.
- Chiu, M., Mason, T.L. and Fink, G.R. (1992) *HTS1* encodes both the cytoplasmic and mitochondrial histidyl-tRNA synthetase of *Saccharomyces cerevisiae*: mutations alter the specificity of compartmentation. *Genetics*, 132, 987-1001.
- Claros, M.G. and Vincens, P. (1996) Computational method to predict mitochondrially imported proteins and their targeting sequences. *FEBS J.*, 241, 779-786.
- Cleary, J.D., Pattamatta, A. and Ranum, L.P.W. (2018) Repeat associated non-ATG (RAN) translation. *J Biol Chem*. pii: jbc.R118.003237.
- Cobucci-Ponzano, B., Rossi, M. and Moracci, M. (2012) Translational recoding in archaea. *Extremophiles*. 16(6):793-803.
- Coots, R.A., Liu, X.M., Mao, Y., Dong, L., Zhou, J., Wan, J., Zhang, X. and Qian, S.B. (2017) m⁶A facilitates eIF4F-independent mRNA translation. *Mol Cell* 68: 504 – 514.e7
- Crawford, R.A. and Pavitt, G.D. (2019) Translational regulation in response to stress in *Saccharomyces cerevisiae*. *Yeast*. 36(1):5-21.
- Crick, F. (1970) Central dogma of molecular biology. *Nature* 227(5258):561-3.
- Cunningham, F., Achuthan, P., Akanni, W., Allen, J., Amode, M.R., Armean, I.M., Bennett, R., Bhai, J., Billis, K., Boddu, S., Cummins, C., Davidson, C., Dodiya, K.J., Gall, A., Girón, C.G., Gil, L., Grego, T., Haggerty, L., Haskell, E., Hourlier, T., Izuogu, O.G., Janacek, S.H., Juettemann, T., Kay, M., Laird, M.R., Lavidas, I., Liu, Z., Loveland, J.E., Marugán, J.C., Maurel, T., McMahon, A.C., Moore, B., Morales, J., Mudge, J.M., Nuhn, M., Ogeh, D., Parker, A., Parton, A., Patricio, M., Abdul Salam, A.I., Schmitt, B.M., Schuilenburg, H., Sheppard, D., Sparrow, H., Stapleton, E., Szuba, M., Taylor, K.,

- Threadgold, G., Thormann, A., Vullo, A., Walts, B., Winterbottom, A., Zadissa, A., Chakiachvili, M., Frankish, A., Hunt, S.E., Kostadima, M., Langridge, N., Martin, F.J., Muffato, M., Perry, E., Ruffier, M., Staines, D.M., Trevanion, S.J., Aken, B.L., Yates, A.D., Zerbino, D.R. and Flicek, P. (2019) Ensembl 2019. *Nucleic Acids Res.* 8.
- Curran, J. and Kolakofsky, D. (1988). Ribosomal initiation from an ACG codon in the Sendai virus P/C mRNA. *EMBO J.* 7, 245–251.
- Dabrowski, M., Bukowy-Bieryllo, Z. and Zietkiewicz, E. (2018) Advances in therapeutic use of a drug-stimulated translational readthrough of premature termination codons. *Mol Med.* 24(1):25.
- de Klerk, E. and 't Hoen, P. A. C. (2015) Alternative mRNA transcription, processing, and translation: insights from RNA sequencing. *Trends Genet. TIG* 31, 128–139.
- Dever, T.E., Dinman, J.D. and Green, R. (2018) Translation Elongation and Recoding in Eukaryotes. *Cold Spring Harb Perspect Biol.* 10(8). pii: a032649.
- Dimmer, K.S., Fritz, S., Fuchs, F., Messerschmitt, M., Weinbach, N., Neupert, W. and Westermann, B. (2002) Genetic basis of mitochondrial function and morphology in *Saccharomyces cerevisiae*. *Mol Biol Cell.* 13(3):847-53.
- Disney, M.D., Liu, B., Yang, W.Y., Sellier, C., Tran, T., Charlet-Berguerand, N. and Childs-Disney, J.L. (2012) A small molecule that targets r(CGG)(exp) and improves defects in fragile X-associated tremor ataxia syndrome. *ACS Chem Biol.* 7(10):1711-8.
- Donovan, J. and Copeland, P.R. (2010) The efficiency of selenocysteine incorporation is regulated by translation initiation factors. *J Mol Biol.* 400(4):659-64.
- D'Souza, A.R. and Minczuk, M. (2018) Mitochondrial transcription and translation: overview. *Essays Biochem.* 62(3):309-320.
- Du, Z., Alekhina, O.M., Vassilenko, K.S. and Simon, A.E. (2017) Concerted action of two 3' cap-independent translation enhancers increases the competitive strength of translated viral genomes. *Nucleic Acids Res.* 45(16):9558-9572.
- Dvinge, H. (2018) Regulation of alternative mRNA splicing: old players and new perspectives. *FEBS Lett.* 592(17):2987-3006.
- El Hage, A., Webb, S., Kerr, A. and Tollervey, D. (2014) Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets intRNA genes, retrotransposons and mitochondria. *PLoS Genet.*, 10, e100471.
- Elfakess, R. and Dikstein, R. (2008) A translation initiation element specific to mRNAs with very short 5'UTR that also regulates transcription. *PLoS One.* 3(8):e3094.
- Emanuelsson, O., Brunak, S., von Heijne, G. and Nielsen, H. (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc.*, 2, 953-971.
- Emmott, E., Jovanovic, M. and Slavov, N. (2019) Ribosome Stoichiometry: From Form to Function. *Trends Biochem Sci.* 44(2):95-109.
- Fatica, A., Dlakić, M. and Tollervey, D. (2002) Naf1p is a box H/ACA snoRNP assembly factor. *RNA.* 8(12):1502-14.
- Fehrenbacher, K.L., Boldogh, I.R. and Pon, L.A. (2005) A role for Jsn1p in recruiting the Arp2/3 complex to mitochondria in budding yeast. *Mol. Biol. Cell,* 16, 5094-5102.
- Fernández-Pevida, A., Kressler, D. and de la Cruz, J. (2015) Processing of preribosomal RNA in *Saccharomyces cerevisiae*. *Wiley Interdiscip. Rev. RNA,* 6, 191-209.
- Ferretti, M.B., Barre, J.L. and Karbstein, K. (2018) Translational Reprogramming Provides a Blueprint for Cellular Adaptation. *Cell Chem Biol.* 25(11):1372-1379.e3.
- Ferretti, M.B., Ghalei, H., Ward, E.A., Potts, E.L. and Karbstein, K. (2017) Rps26 directs mRNA-specific translation by recognition of Kozak sequence elements. *Nat Struct Mol Biol.* 24(9):700-707.
- Firth, A.E. and Brierley, I. (2012) Non-canonical translation in RNA viruses. *J Gen Virol.* 93(Pt 7):1385-409.
- Flohe, L., Günzler, W.A. and Schock, H.H. (1973) Glutathione peroxidase: a selenoenzyme. *FEBS Lett.* 32(1):132-4.
- Fournier, C.T., Cherny, J.J., Truncali, K., Robbins-Pianka, A., Lin, M.S., Krizanc, D. and Weir, M.P. (2012) Amino termini of many yeast proteins map to downstream start codons. *J. Proteome Res.*, 11, 5712-5719.
- Foury, F., Roganti, T., Lecrenier, N. and Purnelle, B. (1998) The complete sequence of the mitochondrial genome of *Saccharomyces cerevisiae*. *FEBS Lett.* 440(3):325-31.

- Fox, T.D and Weiss-Brummer, B. (1980) Leaky +1 and -1 frameshift mutations at the same site in a yeast mitochondrial gene. *Nature*. 288(5786):60-3.
- Friesen, W.J., Johnson, B., Sierra, J., Zhuo, J., Vazirani, P., Xue, X., Tomizawa, Y., Baiazitov, R., Morrill, C., Ren, H., Babu, S., Moon, Y.C., Branstrom, A., Mollin, A., Hedrick, J., Sheedy, J., Elfring, G., Weetall, M., Colacino, J.M., Welch, E.M. and Peltz, S.W. (2018) The minor gentamicin complex component, X2, is a potent premature stop codon readthrough molecule with therapeutic potential. *PLoS One*. 13(10):e0206158.
- Fritsch, C., Herrmann, A., Nothnagel, M., Szafranski, K., Huse, K., Schumann, F., Schreiber, S., Platzer, M., Krawczak, M., Hampe, J. et al. (2012) Genome-wide search for novel human uORFs and N-terminal protein extensions using ribosomal footprinting. *Genome Res.*, 22, 2208-2218.
- Fukasawa, Y., Tsuji, J., Fu, S.C., Tomii, K., Horton, P. and Imai, K. (2015) MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites. *Mol Cell Proteomics*. 14(4):1113-26.
- Fütterer, J., Gordon, K., Sanfaçon, H., Bonneville, J.M. and Hohn, T. (1990) Positive and negative control of translation by the leader sequence of cauliflower mosaic virus pregenomic 35S RNA. *EMBO J*. 9(6):1697-707.
- Gakh, O., Cavadini, P. and Isaya, G. (2002) Mitochondrial processing peptidases. *Biochim. Biophys. Acta*, 1592, 63-77.
- Gelperin, D.M., White, M.A., Wilkinson, M.L., Kon, Y., Kung, L.A., Wise, K.J., Lopez-Hoyo, N., Jiang, L., Piccirillo, S., Yu, H., Gerstein, M., Dumont, M.E., Phizicky, E.M., Snyder, M. and Grayhack, E.J. (2005) Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Dev.*, 19, 2816-2826.
- Gerashchenko, M.V., Lobanov, A.V. and Gladyshev, V.N. (2012) Genomewide ribosome profiling reveals complex translational regulation in response to oxidative stress. *Proc. Natl. Acad. Sci. USA*, 109, 7394-7399.
- Giedroc, D.P. and Cornish, P.V. (2009) Frameshifting RNA pseudoknots: structure and mechanism. *Virus Res*. 139(2):193-208.
- Gomes, R.A., Vicente Miranda, H., Sousa Silva, M., Graça, G., Coelho, A.V., do Nascimento Ferreira, A.E., Cordeiro, C., Freire, A.P. (2008) Protein glycation and methylglyoxal metabolism in yeast: finding peptide needles in protein haystacks. *FEMS Yeast Res*. 8(1):174-81.
- Gonczarowska-Jorge, H., Zahedi, R.P. and Sickmann, A. (2017) The proteome of baker's yeast mitochondria. *Mitochondrion*, 15-21.
- Graf, R., Baum, B. and Braus, G.H. (1993) YMC1, a yeast gene encoding a new putative mitochondrial carrier protein. *Yeast*. 9(3):301-5.
- Gray, M.W., Burger, G. and Lang, B.F. (1999) Mitochondrial evolution. *Science*. 283(5407):1476-81.
- Gross, S. and Moore, C. (2001) Five subunits are required for reconstitution of the cleavage and polyadenylation activities of *Saccharomyces cerevisiae* cleavage factor I. *Proc Natl Acad Sci U S A*. 98(11):6080-5.
- Guenther, U.P., Weinberg, D.E., Zubradt, M.M., Tedeschi, F.A., Stawicki, B.N., Zagore, L.L., Brar, G.A., Licatalosi, D.D., Bartel, D.P., Weissman, J.S. and Jankowsky, E. (2018) The helicase Ded1p controls use of near-cognate translation initiation codons in 5' UTRs. *Nature*, 559, 130-134.
- Gunišová, S., Hronová, V., Mohammad, M.P., Hinnebusch, A.G. and Valášek, L.S. (2018) Please do not recycle! Translation reinitiation in microbes and higher eukaryotes. *FEMS Microbiol Rev*. 42(2):165-192.
- Haimov, O., Sinvani, H. and Dikstein, R. (2015) Cap-dependent, scanning-free translation initiation mechanisms. *Biochim Biophys Acta*. 1849(11):1313-8.
- Hao, B., Gong, W., Ferguson, T.K., James, C.M., Krzycki, J.A. and Chan, M.K. (2002) A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science*. 296(5572):1462-6.
- Hecht, A., Glasgow, J., Jaschke, P.R., Bawazer, L.A., Munson, M.S., Cochran, J.R., Endy, D. and Salit, M. (2017) Measurements of translation initiation from all 64 codons in *E. coli*. *Nucleic Acids Res*. 45(7):3615-3626.
- Heininger, A. (2016) Characterisation of the yeast RNA helicase Prp43. PhD thesis, University of Gottingen, Sweden.
- Helsens, K., Van Damme, P., Degroove, S., Martens, L., Arnesen, T., Vandekerckhove, J. and Gevaert, K. (2011) Bioinformatics analysis of a *Saccharomyces cerevisiae* N-terminal proteome provides evidence

- of alternative translation initiation and post-translational N-terminal acetylation. *J Proteome Res.* 10(8):3578-89.
- Heublein, M. (2014) Mitochondrial Energy Metabolism: Kgd4 is a novel subunit of the α -ketoglutarate dehydrogenase complex. PhD thesis, Department of Biochemistry and Biophysics, Stockholm University.
- Heublein, M., Burguillos, M.A., Vögtle, F.N., Teixeira, P.F., A., I., C., M. and Ott, M. (2014) The novel component Kgd4 recruits the E3 subunit to the mitochondrial α -ketoglutarate dehydrogenase. *Mol. Biol. Cell*, 25, 3342-3349.
- Heublein, M., Ndi, M., Calvo, C.V., Vögtle, F.N. and Ott, M. (2019) Alternative Translation Initiation at a UUG Codon Gives Rise to Two Functional Variants of the Mitochondrial Protein Kgd4. *J Mol Biol.* pii: S0022-2836(19)30100-7.
- Hinnebusch, A.G. (2005) Translational regulation of GCN4 and the general amino acid control of yeast. *Annu Rev Microbiol.* 59:407-50.
- Hinnebusch, A.G., Ivanov, I.P. and Sonenberg N. (2016) Translational control by 5'-untranslated regions of eukaryotic mRNAs. *Science.* 352(6292):1413-6.
- Holcik, M. and Sonenberg, N. (2005) Translational control in stress and apoptosis. *Nat Rev Mol Cell Biol.* 6(4):318-27.
- Holmes, J.B., Akman, G., Wood, S.R., Sakhuja, K., Cerritelli, S.M., Moss, C., Bowmaker, M.R., Jacobs, H.T., Crouch, R.J. and Holt, I.J. (2015) Primer retention owing to the absence of RNase H1 is catastrophic for mitochondrial DNA replication. *Proc. Natl. Acad. Sci. USA*, 112, 9334–9339.
- Hsu, P.Y., Calviello, L., Wu, H.L., Li, F.W., Rothfels, C.J., Ohler, U. and Benfey, P.N. (2016) Superresolution ribosome profiling reveals unannotated translation events in Arabidopsis. *Proc. Natl. Acad. Sci. USA*, 113, E7126-E7135.
- Huang, H.K., Yoon, H., Hannig, E.M. and Donahue, T.F. (1997) GTP hydrolysis controls stringent selection of the AUG start codon during translation initiation in *Saccharomyces cerevisiae*. *GenesDev* 11: 2396–2413.
- Huez, I., Bornes, S., Bresson, D., Creancier, L. and Prats, H. (2001) New vascular endothelial growth factor isoform generated by internal ribosome entry site-driven CUG translation initiation. *Mol. Endocrinol.* 15,2197–2210.
- Humphrey, S.J., James, D.E. and Mann, M. (2015) Protein Phosphorylation: A Major Switch Mechanism for Metabolic Regulation. *Trends Endocrinol Metab.* 26(12):676-687.
- Ingolia, N.T., Brar, G.A., Stern-Ginossar, N., Harris, M.S., Talhouarne, G.J., Jackson, S.E., Wills, M.R. and Weissman, J.S. (2014) Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes. *Cell Rep.*, 8, 1365-1379.
- Ingolia, N.T., Ghaemmhami, S., Newman, J.R. and Weissman, J.S. (2009) Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science*, 324, 218-223.
- Ingolia, N.T., Lareau, L.F. and Weissman, J.S. (2011) Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell.* 147(4):789-802.
- Ivanov, I.P., Firth, A.E., Michel, A.M., Atkins, J.F. and Baranov, P.V. (2011) Identification of evolutionarily conserved non-AUG-initiated N-terminal extensions in human coding sequences. *Nucleic Acids Res.*, 39, 4220-4234.
- Jacks, T. and Varmus, H.E. (1985) Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. *Science.* 230(4731):1237-42.
- Jackson, R.J., Hellen, C.U. and Pestova, T.V. (2010) The mechanism of eukaryotic translation and principles of its regulation. *Nat Rev Mol Cell Biol.* 11(2):113-27.
- Jackson, R.J., Hellen, C.U. and Pestova, T.V. (2012) Termination and post-termination events in eukaryotic translation. *Adv Protein Chem Struct Biol.* 86:45-93.
- James, C.M., Ferguson, T.K., Leykam, J.F. and Krzycki, J.A. (2001) The amber codon in the gene encoding the monomethylamine methyltransferase isolated from *Methanosarcina barkeri* is translated as a sense codon. *J Biol Chem.* 276(36):34252-8.
- Jeong, J.Y., Yim, H.S., Ryu, J.Y., Lee, H.S., Lee, J.H., Seen, D.S. and Kang, S.G. (2012) One-step sequence- and ligation-independent cloning as a rapid and versatile cloning method for functional genomics studies. *Appl. Environ. Microbiol.*, 78, 5440-5443.
- Jia, H., Wang, X., Liu, F., Guenther, U.P., Srinivasan, S., Anderson, J.T. and Jankowsky, E. (2011) The RNA helicase Mtr4p modulates polyadenylation in the TRAMP complex. *Cell.* 145(6):890-901.

- Karnkowska, A., Vacek, V., Zubáčová, Z., Treitli, S.C., Petrželková, R., Eme, L., Novák, L., Žárský, V., Barlow, L.D., Herman, E.K., Soukal, P., Hroudová, M., Doležal, P., Stairs, C.W., Roger, A.J., Eliáš, M., Dacks, J.B., Vlček, Č. and Hampl, V. (2016) A Eukaryote without a Mitochondrial Organelle. *Curr Biol.* 26(10):1274-84.
- Kazak, L., Reyes, A., Duncan, A.L., Rorbach, J., Wood, S.R., Brea-Calvo, G., Gammage, P.A., Robinson, A.J., Minczuk, M. and Holt, I.J. (2013) Alternative translation initiation augments the human mitochondrial proteome. *Nucleic Acids Res.*, 41, 2354-2369.
- Kearse, M.G. and Wilusz, J.E. (2017) Non-AUG translation: a new start for protein synthesis in eukaryotes. *Genes Dev.* 31(17):1717-1731.
- Keren, H., Lev-Maor, G. and Ast, G. (2010) Alternative splicing and evolution: diversification, exon definition and function. *Nat. Rev. Genet.* 11, 345–355.
- Kissová, I., Deffieu, M., Manon, S. and Camougrand, N. (2004) Uth1p is involved in the autophagic degradation of mitochondria. *J Biol Chem.* 279(37):39068-74.
- Klingenberg, M. (1975) Energy transfer in mitochondrial synthesis of ATP; a survey. *Ciba Found Symp.* (31):23-40.
- Koch, A., Gawron, D., Steyaert, S., Ndah, E., Crappé, J., De Keulenaer, S., De Meester, E., Ma, M., Shen, B., Gevaert, K., Van Crielinge, W., Van Damme, P. and Menschaert, G. (2014) A proteogenomics approach integrating proteomics and ribosome profiling increases the efficiency of protein identification and enables the discovery of alternative translation start sites. *Proteomics*, 14, 2688-2698.
- Kochetov, A.V., Sarai, A., Rogozin, I.B., Shumny, V.K. and Kolchanov, N.A. (2005) The role of alternative translation start sites in the generation of human protein diversity. *Mol. Genet. Genomics*, 273, 491-496.
- Kolitz, S.E., Takacs, J.E. and Lorsch, J.R. (2009) Kinetic and thermodynamic analysis of the role of start codon/anticodon base pairing during eukaryotic translation initiation. *RNA.* 15(1):138-52.
- Kosugi, S., Hasebe, M., Tomita, M. and Yanagawa, H. (2009) Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc Natl Acad Sci U S A.* 106(25):10171-6.
- Kozak, M. (1987) An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* 15(20):8125-48.
- Kozak, M. (1990) Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. *Proc Natl Acad Sci U S A.* 87(21):8301-5.
- Kozak, M. (2002) Pushing the limits of the scanning mechanism for initiation of translation. *Gene.* 299(1-2):1-34.
- Kritsiligkou, P., Chatzi, A., Charalampous, G., Mironov, A.J., Grant, C.M. and Tokatlidis, K. (2017) Unconventional targeting of a thiol peroxidase to the mitochondrial intermembrane space facilitates oxidative protein folding. *Cell Rep.*, 18, 2729-2741.
- Kufel, J., Bousquet-Antonelli, C., Beggs, J.D. and Tollervey, D. (2004) Nuclear pre-mRNA decapping and 5' degradation in yeast require the Lsm2-8p complex. *Mol. Cell. Biol.*, 24, 9646-9657.
- Kunze, M. and Berger, J. (2015) The similarity between N-terminal targeting signals for protein import into different organelles and its evolutionary relevance. *Frontiers Physiol.*, 6, 1-27.
- Kurihara, Y., Makita, Y., Kawashima, M., Fujita, T., Iwasaki, S. and Matsui, M. (2018) Transcripts from downstream alternative transcription start sites evade uORF-mediated inhibition of gene expression in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA*, 115, 7831-7836.
- Kwan, T. and Thompson, S.R. (2018) Noncanonical Translation Initiation in Eukaryotes. *Cold Spring Harb Perspect Biol.* pii: a032672.
- Labunskyy, V.M., Hatfield, D.L. and Gladyshev, V.N. (2014) Selenoproteins: molecular pathways and physiological roles. *Physiol Rev.* 94(3):739-77.
- LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A. and Tollervey D. (2005) RNA degradation by the exosome is promoted by a nuclear polyadenylation complex *Cell*, 121(5):713-24
- Lacerda, R., Menezes, J. and Romão, L. (2017) More than just scanning: the importance of cap-independent mRNA translation initiation for cellular stress response and cancer. *Cell Mol Life Sci.* 74(9):1659-1680.
- Lackner, L.L., Ping, H., Graef, M., Murley, A. and Nunnari, J. (2013) Endoplasmic reticulum-associated mitochondria-cortex tether functions in the distribution and inheritance of mitochondria. *Proc Natl Acad Sci U S A.* 110(6):E458-67.

- Lang, B.F., Jakubkova, M., Hegedusova, E., Daoud, R., Forget, L., Brejova, B., Vinar, T., Kosa, P., Fricova, D., Nebohacova, M., Griac, P., Tomaska, L., Burger, G. and Nosek, J. (2014) Massive programmed translational jumping in mitochondria. *Proc Natl Acad Sci U S A*. 111(16):5926-31.
- Lee, A.S., Kranzusch, P.J., Doudna, J.A. and Cate, J.H. (2016) eIF3d is an mRNA cap-binding protein that is required for specialized translation initiation. *Nature* 536: 96 – 99
- Lee, S., Liu, B., Lee, S., Huang, S.-X., B., S. and Qian, S.-B. (2012) Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. *Proc. Natl. Acad. Sci. USA*, 109, E2424–E2432.
- Li, L. and Wang, C.C. (2004) Capped mRNA with a single nucleotide leader is optimally translated in a primitive eukaryote, *Giardia lamblia*. *J Biol Chem*. 279(15):14656-64.
- Liang, H., Chen, X., Yin, Q., Ruan, D., Zhao, X., Zhang, C., McNutt, M.A. and Yin, Y. (2017) PTEN β is an alternatively translated isoform of PTEN that regulates rDNA transcription. *Nat Commun*. 8:14771.
- Liberman, N., Gandin, V., Svitkin, Y.V., David, M., Virgili, G., Jaramillo, M., Holcik, M., Nagar, B., Kimchi, A. and Sonenberg, N. (2015) DAP5 associates with eIF2 β and eIF4AI to promote Internal Ribosome Entry Site driven translation. *Nucleic Acids Res*. 43(7):3764-75.
- Lind, C., Esguerra, M. and Åqvist, J. (2017) A close-up view of codon selection in eukaryotic initiation. *RNA Biol*. 14(7):815-819.
- Llácer, J.L., Hussain, T., Saini, A.K., Nanda, J.S., Kaur, S., Gordiyenko, Y., Kumar, R., Hinnebusch, A.G., Lorsch, J.R. and Ramakrishnan, V (2018) Translational initiation factor eIF5 replaces eIF1 on the 40S ribosomal subunit to promote start-codon recognition. *Elife*. 7. pii: e39273.
- Lobanov, A.V., Heaphy, S.M., Turanov, A.A., Gerashchenko, M.V., Pucciarelli, S., Devaraj, R.R., Xie, F., Petyuk, V.A., Smith, R.D., Klobutcher, L.A., Atkins, J.F., Miceli, C., Hatfield, D.L., Baranov, P.V. and Gladyshev, V.N. (2017) Position-dependent termination and widespread obligatory frameshifting in *Euplotes* translation. *Nat Struct Mol Biol*. 24(1):61-68.
- Lock, P., Ralph, S., Stanley, E., Boulet, I., Ramsay, R. and Dunn, A.R. (1991) Two isoforms of murine hck, generated by utilization of alternative translational initiation codons, exhibit different patterns of subcellular localization. *Mol. Cell. Biol*. 11, 4363–4370.
- Longtin, R. (2004) A forgotten debate: Is selenocysteine the 21st amino acid? *J Natl Cancer Inst*. 96(7):504-5.
- Luke, G.A., Roulston, C., Odon, V., de Felipe, P., Sukhodub, A. and Ryan, M.D. (2013) Lost in translation: The biogenesis of non-LTR retrotransposon proteins. *Mob Genet Elements*. 3(6):e27525.
- Luttermann, C. and Meyers, G. (2014) Two alternative ways of start site selection in human norovirus reinitiation of translation. *J Biol Chem*. 289(17):11739-54.
- Madsen, C.T., Sylvestersen, K.B., Young, C., Larsen, S.C., Poulsen, J.W., Andersen, M.A., Palmqvist, E.A., Hey-Mogensen, M., Jensen, P.B., Treebak, J.T., Lisby, M. and Nielsen, M.L. (2015) Biotin starvation causes mitochondrial protein hyperacetylation and partial rescue by the SIRT3-like deacetylase Hst4p. *Nat Commun*. 6:7726.
- Mailliot, J. and Martin, F. (2018) Viral internal ribosomal entry sites: four classes for one goal. *Wiley Interdiscip Rev RNA*. 9(2).
- Martinez-Salas, E., Francisco-Velilla, R., Fernandez-Chamorro, J. and Embarek, A.M. (2018) Insights into Structural and Mechanistic Features of Viral IRES Elements. *Front Microbiol*. 8:2629.
- Mayes, A.E., Verdone, L., Legrain, P. and Beggs, J.D. (1999) Characterization of Sm-like proteins in yeast and their association with U6 snRNA. *EMBO J*, 18, 4321-4331.
- Meisinger, C., Pfanner, N. and Truscott, K.N. (2006) Isolation of yeast mitochondria. *Methods Mol.Biol.*, 313, 33-39.
- Meisinger, C., Sommer, T. and Pfanner, N. (2000) Purification of *Saccharomyces cerevisiae* mitochondria devoid of microsomal and cytosolic contaminations. *Anal. Biochem.*, 287, 339-342.
- Meng, Z., Jackson, N.L., Shcherbakov, O.D., Choi, H. and Blume, S.W. (2010) The human IGF1R IRES likely operates through a Shine-Dalgarno-like interaction with the G961 loop (E-site) of the 18S rRNA and is kinetically modulated by a naturally polymorphic polyU loop. *J Cell Biochem*. 110(2):531-44.
- Menschaert, G., Van Crielinge, W., Notelaers, T., Koch, A., Crappé, J., Gevaert, K. and Van Damme, P. (2013) Deep proteome coverage based on ribosome profiling aids mass spectrometrybased protein and peptide discovery and provides evidence of alternative translation products and near-cognate translation initiation events. *Mol. Cell. Proteomics*, 12, 1780-1790.

- Merz, S. and Westermann, B. (2009) Genome-wide deletion mutant analysis reveals genes required for respiratory growth, mitochondrial genome maintenance and mitochondrial protein synthesis in *Saccharomyces cerevisiae*. *Genome Biol.* 10(9):R95.
- Mewes, H.W., Albermann, K., Bähr, M., Frishman, D., Gleissner, A., Hani, J., Heumann, K., Kleine, K., Maierl, A., Oliver, S.G., Pfeiffer, F. and Zollner, A. (1997) Overview of the yeast genome. *Nature*. 387(6632 Suppl):7-65.
- Meyer, K.D. (2018) m⁶A-mediated translation regulation. *Biochim Biophys Acta Gene Regul Mech.* pii: S1874-9399(18)30105-6.
- Meyer, K.D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C.E. and Jaffrey, S.R. (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell*. 2012 149(7):1635-46.
- Michel, A.M., Andreev, D.E. and Baranov, P.V. (2014) Computational approach for calculating the probability of eukaryotic translation initiation from ribo-seq data that takes into account leaky scanning. *BMC Bioinformatics*. 15:380.
- Miras, M., Rodríguez-Hernández, A.M., Romero-López, C., Berzal-Herranz, A., Colchero, J., Aranda, M.A. and Truniger, V. (2018) A Dual Interaction Between the 5'- and 3'-Ends of the Melon Necrotic Spot Virus (MNSV) RNA Genome Is Required for Efficient Cap-Independent Translation. *Front Plant Sci.* 9:625.
- Mitchell, S.F. and Parker, R. (2015) Modifications on translation initiation. *Cell* 163: 796 – 798
- Monteuuis, G., Miścicka, A., Świrski, M., Zenad, L., Niemitalo, O., Wrobel, L., Alam, J., Chacinska, A., Kastaniotis, A.J. and Kufel, J. (2019) Non-canonical translation initiation in yeast generates a cryptic pool of mitochondrial proteins. *Nucleic Acid Research*, doi: 10.1093/nar/gkz301
- Monteuuis, G., Suomi, F., Kerätär, J.M., Masud, A.J. and Kastaniotis, A.J. (2017) A conserved mammalian mitochondrial isoform of acetyl-CoA carboxylase ACC1 provides the malonyl-CoA essential for mitochondrial biogenesis in tandem with ACSF3. *Biochem J.* 474(22):3783-3797.
- Morgenstern, M., Stiller, S.B., Lübbert, P., Peikert, C.D., Dannenmaier, S., Drepper, F., Weill, U., Höß, P., Feuerstein, R., Gebert, M., Bohnert, M., van der Laan, M., Schuldiner, M., Schütze, C., Oeljeklaus, S., Pfanner, N., Wiedemann, N. and Warscheid, B. (2017) Definition of a high-confidence mitochondrial proteome at quantitative scale. *Cell Rep.*, 19, 2836–2852.
- Mosallanejad, K., Sekine, Y., Ishikura-Kinoshita, S., Kumagai, K., Nagano, T., Matsuzawa, A., Takeda, K., Naguro, I. and Ichijo, H. (2014) The DEAH-box RNA helicase DHX15 activates NF-κB and MAPK signaling downstream of MAVS during antiviral responses. *Sci. Signal.*, 7, ra40.
- Mouilleron, H., Delcourt, V. and Roucou, X. (2016) Death of a dogma: eukaryotic mRNAs can code for more than one protein. *Nucleic Acids Res.*, 44, 14-23.
- Mumberg, D., Müller, R. and Funk, M. (1995) Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene*, 156, 119-122.
- Na, C.H., Barbhuiya, M.A., Kim, M.S., Verbruggen, S., Eacker, S.M., Pletnikova, O., Troncoso, J.C., Halushka, M.K., Menschaert, G., Overall, C.M. and Pandey, A. (2018) Discovery of non-canonical translation initiation sites through mass spectrometric analysis of protein N-termini. *Genome Res.* 28(1):25-36.
- Nakai, K. and Horton, P. (1999) PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem Sci.* 24(1):34-6.
- Natsoulis, G., Hilger, F. and Fink, G.R. (1986) The *HTS1* gene encodes both the cytoplasmic and mitochondrial histidine tRNA synthetases of *S. cerevisiae*. *Cell*, 46, 235-243.
- Nosek, J., Tomaska, L., Burger, G. and Lang, B.F. (2015) Programmed translational bypassing elements in mitochondria: structure, mobility, and evolutionary origin. *Trends Genet.* 31(4):187-94.
- Oyama, M., Kozuka-Hata, H., Suzuki, Y., Semba, K., Yamamoto, T. and Sugano, S. (2007) Diversity of translation start sites may define increased complexity of the human short ORFeome. *Mol. Cell. Proteomics*, 6, 1000-1006.
- Pakos-Zebrucka, K., Koryga, I., Mnich, K., Ljubic, M., Samali, A. and Gorman, A.M. (2016) The integrated stress response. *EMBO Rep.*, 17, 1374-1395.
- Palade, G.E. (1955) A small particulate component of the cytoplasm. *J Biophys Biochem Cytol.* 1(1):59-68.
- Pavitt, G.D. and Ron, D. (2012) New insights into translational regulation in the endoplasmic reticulum unfolded protein response. *Cold Spring Harb Perspect Biol.* 4(6). pii: a012278.

- Peabody, D.S., (1989) Translation initiation at non-AUG triplets in mammalian cells. *J. Biol. Chem.* 264, 5031–5035.
- Pedrajas, J.R., Kosmidou, E., Miranda-Vizuete, A., Gustafsson, J.A., Wright, A.P. and Spyrou, G. (1999) Identification and functional characterization of a novel mitochondrial thioredoxin system in *Saccharomyces cerevisiae*. *J Biol Chem.* 274(10):6366-73.
- Pelechano, V., Wei, W. and Steinmetz, L.M. (2013) Extensive transcriptional heterogeneity revealed by isoform profiling. *Nature*, 497, 127-131.
- Pereira, K.D., Tamborlin, L., Meneguello, L., de Proença, A.R., Almeida, I.C., Lourenço, R.F. and Luchessi, A.D. (2016) Alternative Start Codon Connects eIF5A to Mitochondria. *J Cell Physiol.* 231(12):2682-9.
- Plaza, S., Menschaert, G. and Payre, F. (2017) In search of lost small peptides. *Annu. Rev. Cell. Dev. Biol.*, 33, 391-416.
- Pooggin, M.M. and Ryabova, L.A. (2018) Ribosome Shunting, Polycistronic Translation, and Evasion of Antiviral Defenses in Plant Pararetroviruses and Beyond. *Front Microbiol.* 9:644.
- Prats, A.C., De Billy, G., Wang, P. and Darlix, J.L. (1989) CUG initiation codon used for the synthesis of a cell surface antigen coded by the murine leukemia virus. *J. Mol. Biol.* 205, 363–372.
- Pueyo, J.I., Magny, E.G. and Couso, J.P. (2016) New peptides under the s(ORF)ace of the genome. *Trends Biochem. Sci.*, 41, 665-678.
- Ramakrishnan, V. (2002) Ribosome structure and the mechanism of translation. *Cell.* 108(4):557-72.
- Ramakrishnan, V. (2014) The ribosome emerges from a black box. *Cell.* 159(5):979-984.
- Reuter, K., Biehl, A., Koch, L. and Helms, V. (2016) PreTIS: A tool to predict non-canonical 5' UTR translational initiation sites in human and mouse. *PLoSComput. Biol.*, 12, e1005170.
- Rojas-Duran, M.F. and Gilbert, W.V. (2012) Alternative transcription start site selection leads to large differences in translation activity in yeast. *RNA*, 18, 2299-2305.
- Rossmannith, W. (2011) Localization of human RNase Z isoforms: dual nuclear/mitochondrial targeting of the ELAC2 gene product by alternative translation initiation. *PLoS One*, 6, e19152.
- Rother, M. and Krzycki, J.A. (2010) Selenocysteine, pyrrolysine, and the unique energy metabolism of methanogenic archaea. *Archaea*. pii: 453642.
- Rouillard, J.M., Brendolise, C. and Lacroute, F. (2000) Rna14p, a component of the yeast nuclear cleavage/polyadenylation factor I, is also localised in mitochondria. *Mol Gen Genet.* 262(6):1103-12.
- Sambrook, J., Fritsch, EF. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory.
- Samir, P., Browne, C.M., Rahul, Sun, M., Shen, B., Li, W., Frank, J. and Link, A.J. (2018) Identification of Changing Ribosome Protein Compositions using Mass Spectrometry. *Proteomics* 18(20):e1800217.
- Segev, N. and Gerst, J.E. (2018) Specialized ribosomes and specific ribosomal protein paralogs control translation of mitochondrial proteins. *J Cell Biol.* 217(1):117-126.
- Sendoel, A., Dunn, J.G., Rodriguez, E.H., Naik, S., Gomez, N.C., Hurwitz, B., Levorse, J., Dill, B.D., Schramek, D., Molina, H. et al. (2017) Translation from unconventional 5' start sites drives tumour initiation. *Nature*, 541, 494-499.
- Sharma, P., Yan, F., Doronina, V.A., Escuin-Ordinas, H., Ryan, M.D. and Brown, J.D. (2012) 2A peptides provide distinct solutions to driving stop-carry on translational recoding. *Nucleic Acids Res.* 40(7):3143-51.
- Shatsky, I.N., Terenin, I.M., Smirnova, V.V. and Andreev, D.E. (2018). Cap-Independent Translation: What's in a Name? *Trends Biochem Sci.* 43(11):882-895.
- Shirokikh, N.E. and Preiss, T. (2018) Translation initiation by cap-dependent ribosome recruitment: Recent insights and open questions. *Wiley Interdiscip Rev RNA.* 9(4):e1473.
- Skowronek, E., Grzechnik, P., Späth, B., Marchfelder, A. and Kufel, J. (2014) tRNA 3' processing in yeast involves tRNase Z, Rex1, and Rrp6. *RNA*, 20, 115-130.
- Slonimski, P.P., and Ephrussi, B. (1949). Action de l'acriflavine sur les levures. V. Le système des cytochromes des mutants "petite colonie." *Ann. Inst. Pasteur* 77, 47-63.
- Sokol, A.M., Sztolszterer, M.E., Wasilewski, M., Heinz, E. and Chacinska, A. (2014) Mitochondrial protein translocases for survival and wellbeing. *FEBS Lett.*, 588, 2484-2495.

- Spealman, P., Naik, A.W., May, G.E., Kuersten, S., Freeberg, L., Murphy, R.F. and McManus, J. (2018) Conserved non-AUG uORFs revealed by a novel regression analysis of ribosome profiling data. *Genome Res*, 28.
- Spriggs, K.A., Bushell, M. and Willis, A.E. (2010) Translational regulation of gene expression during conditions of cell stress. *Mol Cell*. 40(2):228-37.
- Spriggs, K.A., Cobbold, L.C., Jopling, C.L., Cooper, R.E., Wilson, L.A., Stoneley, M., Coldwell, M.J., Poncet, D., Shen, Y.C., Morley, S.J., Bushell, M., Willis, A.E. (2009) Canonical initiation factor requirements of the Myc family of internal ribosome entry segments. *Mol Cell Biol*. 29(6):1565-74.
- Sriram, A., Bohlen, J. and Teleman, A.A. (2018) Translation acrobatics: how cancer cells exploit alternate modes of translational initiation. *EMBO Rep*. 19(10). pii: e45947.
- Starck, S.R., Jiang, V., Pavon-Eternod, M., Prasad, S., McCarthy, B., Pan, T. and Shastri, N. (2012) Leucine-tRNA initiates at CUG start codons for protein synthesis and presentation by MHC class I. *Science*. 336(6089):1719-23.
- Starck, S.R., Tsai, J.C., Chen, K., Shodiya, M., Wang, L., Yahiro, K., Martins-Green, M., Shastri, N. and Walter, P. (2016) Translation from the 5' untranslated region shapes the integrated stress response. *Science*, 351, aad3867.
- Stram, A.R. and Payne, R.M. (2016) Post-translational modifications in mitochondria: protein signaling in the powerhouse. *Cell. Mol. Life Sci.*, 73, 4063-4073.
- Suomi, F., Menger, K.E., Monteuuis, G., Naumann, U., Kursu, V.A., Shvetsova, A. and Kastaniotis, A.J. (2014) Expression and evolution of the non-canonically translated yeast mitochondrial acetyl-CoA carboxylase Hfa1p. *PLoS One*, 9, e114738.
- Tang, H.L., Yeh, L.S., Chen, N.K., Ripmaster, T., Schimmel, P. and Wang, C.C. (2004) Translation of a yeast mitochondrial tRNA synthetase initiated at redundant non-AUG codons. *J. Biol. Chem.*, 279, 49656-49663.
- Tate, W.P., Cridge, A.G. and Brown, C.M. (2018) 'Stop' in protein synthesis is modulated with exquisite subtlety by an extended RNA translation signal. *Biochem Soc Trans*. pii: BST20180190.
- Teixeira, P.F. and Glaser, E. (2013) Processing peptidases in mitochondria and chloroplasts. *Biochim Biophys Acta*. 1833(2):360-70.
- Thakur, A. and Hinnebusch, A.G. (2018) eIF1 Loop 2 interactions with Met-tRNA_i control the accuracy of start codon selection by the scanning preinitiation complex. *Proc Natl Acad Sci U S A*. 115(18):E4159-E4168.
- Tharun, S. (2009) Roles of eukaryotic Lsm proteins in the regulation of mRNA function. *Int Rev Cell Mol Biol*. 272:149-89.
- Todisco, S., Agrimi, G., Castegna, A. and Palmieri, F. (2006) Identification of the mitochondrial NAD⁺ transporter in *Saccharomyces cerevisiae*. *J Biol Chem*. 281(3):1524-31.
- Torchet, C., Bousquet-Antonelli, C., Milligan, L., Thompson, E., Kufel, J. and Tollervey, D. (2002) Processing of 3' extended read-through transcripts by the exosome can generate functional mRNAs. *Mol. Cell*, 9, 1285-1296.
- Touriol, C., Bornes, S., Bonnal, S., Audigier, S., Prats, H., Prats, A.C. and Vagner, S. (2003) Generation of protein isoform diversity by alternative initiation of translation at non-AUG codons. *Biol. Cell*, 95, 169-178.
- Trotter, E.W. and Grant, C.M. (2005) Overlapping roles of the cytoplasmic and mitochondrial redox regulatory systems in the yeast *Saccharomyces cerevisiae*. *Eukaryot Cell*. 4(2):392-400.
- Truniger, V., Miras, M., Aranda, M.A. (2017) Structural and Functional Diversity of Plant Virus 3'-Cap-Independent Translation Enhancers (3'-CITEs). *Front Plant Sci*. 8:2047.
- Van Damme, P., Gawron, D., Van Crielinge, W. and Menschaert, G. (2014) N-terminal proteomics and ribosome profiling provide a comprehensive view of the alternative translation initiation landscape in mice and men. *Mol. Cell. Proteomics*, 13, 1245-1261.
- van der Horst, S., Snel, B., Hanson, J.3. and Smeekens, S. (2019) Novel pipeline identifies new upstream ORFs and non-AUG initiating main ORFs with conserved amino acid sequences in the 5' leader of mRNAs in *Arabidopsis thaliana*. *RNA*. 25(3):292-304.
- van Roermund, C.W., Htetema, E.H., van den Berg, M., Tabak, H.F. and Wanders, R.J. (1999) Molecular characterization of carnitine-dependent transport of acetyl-CoA from peroxisomes to mitochondria

- in *Saccharomyces cerevisiae* and identification of a plasma membrane carnitine transporter, Agp2p. EMBO J. 1999 Nov 1;18(21):5843-52.
- Vanáčová, S., Wolf, J., Martin, G., Blank, D., Dettwiler, S., Friedlein, A., Langen, H., Keith, G. and Keller, W. (2005) A new yeast poly(A) polymerase complex involved in RNA quality control. PLoS Biol. 3(6):e189.
- Vasiljeva, L., Kim, M., Mutschler, H., Buratowski, S. and Meinhart, A. (2008) The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. Nat Struct Mol Biol. 15(8):795-804.
- Vögtle, F.N., Burkhart, J.M., Rao, S., Gerbeth, C., Hinrichs, J., Martinou, J.C., Chacinska, A., Sickmann, A., Zahedi, R.P. and Meisinger, C. (2012) Intermembrane space proteome of yeast mitochondria. Mol. Cell. Proteomics, 11, 1840-1852.
- Vögtle, F.N., Wortelkamp, S., Zahedi, R.P., Becker, D., Leidhold, C., Gevaert, K., Kellermann, J., Voos, W., Sickmann, A., Pfanner, N. and Meisinger, C. (2009) Global analysis of the mitochondrial N-proteome identifies a processing peptidase critical for protein stability. Cell. 139(2):428-39.
- von Heijne, G., Steppuhn, J. and Herrmann, R.G. (1989) Domain structure of mitochondrial and chloroplast targeting peptides. Eur. J. Biochem., 180, 535-545.
- Walsh, D., Mathews, M.B. and Mohr, I. (2013) Tinkering with translation: protein synthesis in virus-infected cells. Cold Spring Harb Perspect Biol. 5(1):a012351.
- Wamboldt, Y., Mohammed, S., Elowsky, C., Wittgren, C., de Paula, W.B. and Mackenzie, S.A. (2009) Participation of leaky ribosome scanning in protein dual targeting by alternative translation initiation in higher plants. Plant Cell, 21, 157-167.
- Wang, X., Hou, J., Quedenau, C. and Chen, W. (2016) Pervasive isoform-specific translational regulation via alternative transcription start sites in mammals. Mol. Syst. Biol., 12, 875.
- Weingarten-Gabbay, S., Elias-Kirma, S., Nir, R., Gritsenko, A.A., Stern-Ginossar, N., Yakhini, Z., Weinberger, A. and Segal, E. (2016) Comparative genetics. Systematic discovery of cap-independent translation sequences in human and viral genomes. Science 351: aad4939
- Weiss, R.B., Huang, W.M. and Dunn, D.M. (1990) A nascent peptide is required for ribosomal bypass of the coding gap in bacteriophage T4 gene 60. Cell. 62(1):117-26.
- Wiedemann, N. and Pfanner, N. (2017) Mitochondrial Machineries for Protein Import and Assembly. Annu Rev Biochem. 86:685-714.
- Wojciechowska, M., Olejniczak, M., Galka-Marciniak, P., Jazurek, M. and Krzyzosiak, W.J. (2014) RAN translation and frameshifting as translational challenges at simple repeats of human neurodegenerative disorders. Nucleic Acids Res. 42(19):11849-64.
- Wu, B. and White, K.A. (1999) A primary determinant of cap-independent translation is located in the 3'-proximal region of the tomato bushy stunt virus genome. J Virol. 73(11):8982-8.
- Wu, J., Huang, H.Y. and Hopper, A.K. (2013) A rapid and sensitive non-radioactive method applicable for genome-wide analysis of *Saccharomyces cerevisiae* genes involved in small RNA biology. Yeast. 30(4):119-28.
- Xue, H. and Wong, J.T. (2017) Future of the Genetic Code. Life (Basel). 7(1). pii: E10.
- Yang, M., Jensen, R.E., Yaffe, M.P., Oppliger, W. and Schatz, G. (1988) Import of proteins into yeast mitochondria: the purified matrix processing protease contains two subunits which are encoded by the nuclear MAS1 and MAS2 genes. EMBO J. 7(12):3857-62.
- Yang, P.K., Hoareau, C., Froment, C., Monsarrat, B., Henry, Y. and Chanfreau, G. (2005) Cotranscriptional recruitment of the pseudouridylsynthetase Cbf5p and of the RNA binding protein Naf1p during H/ACA snoRNP assembly. Mol Cell Biol. 25(8):3295-304.
- Yang, Y., Fan, X., Mao, M., Song, X., Wu, P., Zhang, Y., Jin, Y., Yang, Y., Chen, L.L., Wang, Y., Wong, C.C., Xiao, X. and Wang, Z. (2017) Extensive translation of circular RNAs driven by N6-methyladenosine. Cell Res. 27(5):626-641.
- Yen, W.L., Legakis, J.E., Nair, U. and Klionsky, D.J. (2007) Atg27 is required for autophagy-dependent cycling of Atg9. Mol Biol Cell. 18(2):581-93.
- Yeom, J., Ju, S., Choi, Y., Paek, E. and Lee, C. (2017) Comprehensive analysis of human protein N-termini enables assessment of various protein forms. Sci. Rep., 7, 6599.

- Yoshihisa, T., Yunoki-Esaki, K., Ohshima, C., Tanaka, N., Endo, T. (2003) Possibility of cytoplasmic pre-tRNA splicing: the yeast tRNA splicing endonuclease mainly localizes on the mitochondria. *Mol Biol Cell*. 14(8):3266-79.
- Young, E.T. and Pilgrim, D. (1985) Isolation and DNA sequence of ADH3, a nuclear gene encoding the mitochondrial isozyme of alcohol dehydrogenase in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, 3024-3034.
- Zahedi, R.P., Sickmann, A., Boehm, A.M., Winkler, C., Zufall, N., Schönfisch, B., Guiard, B., Pfanner, N. and Meisinger, C. (2006) Proteomic analysis of the yeast mitochondrial outer membrane reveals accumulation of a subclass of preproteins. *Mol Biol Cell*. 17(3):1436-50.
- Zhang, F. and Hinnebusch, A.G. (2011) An upstream ORF with non-AUG start codon is translated *in vivo* but dispensable for translational control of GCN4 mRNA. *Nucleic Acids Res*. 39(8):3128-40.
- Zhang, P., He, D., Xu, Y., Hou, J., Pan, B.F., Wang, Y., Liu, T., Davis, C.M., Ehli, E.A., Tan, L., Zhou, F., Hu, J., Yu, Y., Chen, X., Nguyen, T.M., Rosen, J.M., Hawke, D.H., Ji, Z. and Chen, Y. (2017) Genome-wide identification and differential analysis of translational initiation. *Nat. Commun.*, 8, 1749.
- Zhang, Y., Qi, H., Taylor, R., Xu, W., Liu, L.F. and Jin, S. (2007) The role of autophagy in mitochondria maintenance: characterization of mitochondrial functions in autophagy-deficient *S. cerevisiae* strains. *Autophagy*, 3, 337-346.
- Zhou, J., Wan, J., Shu, X.E., Mao, Y., Liu, X.M., Yuan, X., Zhang, X., Hess, M.E., Brüning, J.C. and Qian, S.B. (2018) N⁶-Methyladenosine Guides mRNA Alternative Translation during Integrated Stress Response. *Mol Cell*. 69(4):636-647.e7.
- Zinoviev, A., Hellen, C.U.T. and Pestova, T.V. (2015) Multiple mechanisms of reinitiation on bicistronic calicivirus mRNAs. *Mol Cell*. 57(6):1059-1073.